5-2016

Refining Methods for Quantifying Macroinvertebrate Estimates of Preference for Use in Stream Bioassessments

Ellen F. Wakeley Tomlinson
Utah State University

REFINING METHODS FOR QUANTIFYING MACROINVERTEBRATE ESTIMATES OF PREFERENCE FOR USE IN STREAM BIOASSESSMENTS

by

Ellen F. Wakeley Tomlinson

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Ecology

Approved:

_____	_____
Charles P. Hawkins Major Professor	D. Richard Cutler Committee Member
_____	_____
Nancy O. Mesner Committee Member	Mark R. McLellan Vice President for Research and Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2016

ABSTRACT

Refining Methods for Quantifying Macroinvertebrate Estimates of Preference for Use in Stream Bioassessments

by

Ellen F. Wakeley Tomlinson, Master of Science

Utah State University, 2016

Major Professor: Dr. Charles P. Hawkins
Department: Watershed Sciences

Two-thirds of United States stream length is in either fair or poor biological condition. However, we do not yet have reliable, quantitative tools to diagnose site-specific causes of biological impairment. One way to diagnose causes of impairment is to compare the environmental tolerances or preferences of the taxa expected at a site to those of the observed community. Stream ecologists have derived tolerance values (TVs) from field data for use in causal analysis, but inconsistencies across studies cast doubt on the accuracy of these TVs. Published TVs may not agree with one another for several reasons, including: differences in the methods used to estimate TVs, confounding among environmental variables, effects of environmental interactions on taxon abundance, and bias from incomplete sampling of a taxon's niche space. My objectives were to determine if and how these four potential problems affect TVs and to determine if TVs can be used reliably in causal assessments. I collected macroinvertebrate, substrate, temperature, conductivity, and velocity data from 45 streams in the western U.S.

Analyses of data from this field study suggest that incomplete sampling can profoundly influence TVs, whereas interactions may have little effect. Significant spatial correlations between environmental variables suggest that confounding may also affect the derivation of TVs, but the magnitude of this effect is unclear. Also, though each method used to estimate TVs has limitations, TVs derived from different methods appear to reveal the same environmental signal, and there may be little reason to prefer one method over another in a causal analysis. Over last few decades, researchers have begun to add biological traits, including estimates of preference and tolerance, to the suite of metrics used in bioassessments. As descriptors of important biological traits, TVs should help water quality managers reliably diagnose and combat the causes of biological stream impairment.

(110 pages)

PUBLIC ABSTRACT

Refining Methods for Quantifying Macroinvertebrate Estimates of Preference for Use in Stream Bioassessments

Ellen F. Wakeley Tomlinson

Tracking the health of streams is increasingly important as people place ever-greater demands on the Nation's water supply. Surveys of stream biological communities – insects and other macroinvertebrates – indicate that two-thirds of the stream length in the U.S. is in fair to poor condition. Defining and monitoring stream health through bioassessments requires a method that pinpoints what specific stressors are causing poor stream condition, so managers can better target restoration or remediation activities. To meet this requirement, some stream ecologists have calculated taxa- and stressor-specific tolerance values (TVs) to compare environmental tolerances of the taxa expected at a site with those of the observed community. However, there are inconsistencies among TVs either calculated by different methods or derived from different geographic regions. The objective of this field and statistical study was to determine if inherent biases of field sampling, such as interactive effects and regional sampling, and different methods of calculating TVs can influence the reliability of TVs as diagnostic tools. The results indicate that some problems inherent in field sampling, especially incomplete sampling, can influence TVs. However, despite documented differences in TVs among studies, they appear sensitive enough to identify patterns in species composition and abundance

influenced by different major stressors. The utility of TVs in stream bioassessments should increase as estimation of TVs continues to improve.

ACKNOWLEDGMENTS

I am grateful for the support of the S.J. & Jessie E. Quinney Foundation throughout my field research and to the Quinney Fellowship Program for support of my studies while I was in residence at USU. The PIBO (PACFISH/INFISH Biological Opinion) Program of the U.S. Forest Service made my field sampling possible. In the summer of 2010, the PIBO program provided transportation to many of my sampling sites, meals, and companionship for weeks at a time, and I am very grateful for their help. I would also like to thank the BLM/USU National Aquatic Monitoring Center (or "Bug Lab"), whose members sorted and identified all the macroinvertebrates in my field samples.

Several USU faculty and staff members and friends were instrumental in helping me complete this thesis, and I am forever grateful for their support. My lab mates – John Olson, Jake Vander Laan, Ryan Hill, and Nora Burbank Washbourne – helped in a variety of ways, including field sampling, map creation, assistance with statistics, and moral support. Susan Durham and Richard Cutler guided my statistical analyses. Jim MacMahon provided guidance about the direction my thesis work should take early in the process. Eric Hazelton, Mark Minton, Ashley D'Antonio, and Kelly Goonan provided assistance with statistics, short-term housing, and friendly encouragement. Brian Bailey kept me and my paperwork organized and moved this thesis through the publication process.

I would also like to thank my advisor, Chuck Hawkins, and my committee members, Nancy Mesner and Richard Cutler, for their questions, edits, suggestions, and

guidance. Their contributions not only improved this manuscript, but expanded my understanding of research and of environmental data analysis. I am grateful for the opportunity to conduct research on an original and challenging topic, and for their support throughout the process.

Last but not least, I'd like to thank my family. My parents – both of whom completed graduate degree programs at USU – provided endless scientific, editorial, financial, and moral support, and I would not have completed this thesis without them. I would also like to thank my husband, Steve, for reminding me what it was like to be a student and for keeping our lives and home going while I was engrossed in tolerance values.

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND DEFINITION OF THE PROBLEM

Running waters are one of the most threatened types of ecosystems on the planet, and the biological integrity of streams and rivers is being increasingly threatened by human actions (Malmqvist and Rundle 2002, Allan 2004). Many streams in the United States do not meet water quality objectives set by the 1972 Clean Water Act (33 U.S.C. 1344), whose goal is to "restore and maintain the chemical, physical, and biological integrity of the Nation's waters". Under section 303(d) of the Clean Water Act, when stream water quality standards are not met, a list of those streams must be submitted to the Environmental Protection Agency. States must then develop a TMDL (Total Maximum Daily Load) for each non-compliant pollutant. A TMDL is the maximum amount of a pollutant that a stream can accommodate and still meet water quality standards. If required, biological assessments (bioassessments) can then be used to determine if pollution control actions taken as a result of the TMDL are restoring the biological integrity of the stream (USEPA 1994).

Bioassessments require establishing a reference condition and developing a way of measuring biological condition. The reference condition is a benchmark to which the condition of the water body of interest is compared (Stoddard et al. 2006, USEPA 2006, Hawkins et al. 2010). Reference condition can refer to several different biological states, including historical condition, least disturbed condition, minimally disturbed condition, and best attainable condition (Stoddard et al. 2006). Once a benchmark condition is established, samples of macroinvertebrates or other organisms are collected from the

water body of interest. Data on the types and number of taxa collected are then typically summarized into an index whose value reflects the overall biological condition of the water body (USEPA 2006).

Bioassessments can indicate whether a water body is biologically impaired but cannot reveal what stressors are causing the impairment (Norton et al. 2000). One way to potentially diagnose the cause of biological impairment is to compare the specific environmental preferences of the taxa observed at a site to the preferences of the expected taxa (Meador et al. 2008, Turley et al. 2014). For example, the mean preferences for water temperature and percent fine sediment of taxa expected at a site of interest can be compared to the mean preferences of taxa observed at the site (Table 1.1). In this hypothetical example, the mean temperature preferences for the observed and expected taxa are very similar (10.8˚ and 10.4˚C, respectively), which would imply that water temperature is not the cause of impairment. In contrast, the mean percent fine sediment preferences are more dissimilar (31% for the expected community and 58% for the observed community). This difference suggests that the observed taxa prefer or tolerate a higher percentage of fine sediment than the expected taxa and, therefore, an increase in fine sediment may be the cause of biological impairment.

Several authors have estimated macroinvertebrate preferences/tolerances for various environmental variables from survey data. These environmental variables include temperature (Brandt 2001, Yuan 2006, Huff et al. 2008), fine sediment in the streambed (Yuan 2006, Carlisle et al. 2007, Huff et al. 2008, Bryce et al. 2010, Relyea et al. 2012, Hubler et al. 2016), nutrients (Carlisle et al. 2007, Meador et al. 2008), salinity

(Horrigan et al. 2007), and stream water acidity (Tripole et al. 2008). Estimates have also been generated for general pollution gradients (Lenat 1993, Bressler et al. 2006) and correlated sets of variables, such as temperature and dissolved oxygen (Carlisle et al. 2007). The data used to estimate these preferences were collected over a variety of spatial scales, including single, 300-km^2 watersheds (Tripole et al. 2008), individual states (Lenat 1993, Brandt 2001, Bressler et al. 2006, Huff et al. 2008, Hubler et al. 2016), multiple states (Meador et al. 2008, Relyea et al. 2012), and very large portions of the country (Yuan 2006, Carlisle et al. 2007, Bryce et al. 2010).

Estimates of preference or tolerance can be calculated from survey data in a variety of ways. Often they are calculated as averages of the environmental variables across the sampled sites weighted by the abundances of species at each site (e.g., Brandt 2001, Carlisle et al. 2007, Huff et al. 2008, and Bryce et al. 2010, Hubler et al. 2016). A weighted average is interpreted as identifying a taxon's preferred or optimum level of some environmental variable (ter Braak and Barendregt 1986, Huff et al. 2008). Macroinvertebrate preferences can also be described with parametric and nonparametric regressions (Yuan 2006). Alternatively, tolerance limits can be estimated (Yuan 2006, Relyea et al. 2012). Tolerance limits can be calculated as abundance-weighted standard deviations (Brandt 2001, Huff et al. 2008) or by arbitrarily specifying thresholds along an environmental gradient below or above which some percentage of the individuals in a taxon occur (e.g., Relyea et al. 2012).

Definition of the problem

Using invertebrate samples to diagnose the causes of biological impairment in streams requires reliable and repeatable estimates of environmental preferences for various taxa. However, inconsistencies in these estimates among different authors cast doubt on their general accuracy. For example, comparing Spearman's rank correlation coefficients between taxon-specific tolerance values for percent fine sediment estimated in five studies (Yuan 2006, Carlisle et al. 2007, Huff et al. 2008, Bryce et al. 2010, and Relyea et al. 2012) revealed a range of correlation values from 0.24 to 0.68 (Table 1.2). Only two of these correlations were significant with a Bonferroni correction for multiple comparisons: Huff et al. 2008 and Bryce et al. 2010 (0.68), and Yuan 2006 and Carlisle et al. 2007 (0.63). These relatively low correlation coefficients indicate either severe imprecision in, or outright disagreement among, estimates, which implies that many or all of them are unreliable.

Published estimates of preference or tolerance may differ for many reasons. Different analytical methods might affect estimates. Correlations among environmental variables in the data used to estimate preferences may confound attribution of how taxa respond to specific stressors. Interactions among environmental variables may influence taxon presence and abundance and, therefore, estimates of preference. Incomplete sampling of a taxon's niche space associated with small-scale field surveys may result in biased estimates of preference. In addition, estimates of preference generated from habitat-use data alone may be biased if habitat availability is not incorporated into calculations. The research described in this thesis was designed to determine if and how

these potential sources of error influence estimates of habitat preference for 63 western macroinvertebrate taxa.

REFERENCES

Allan, J. D. 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. Annual Review of Ecology, Evolution, and Systematics 35:257-284.

Brandt, D. 2001. Temperature preferences and tolerances for 137 common Idaho macroinvertebrate taxa. Idaho Department of Environmental Quality. Coeur d'Alene, Idaho.

Bressler, D. W., J. B. Stribling, M. J. Paul, and M. B. Hicks. 2006. Stressor tolerance values for benthic macroinvertebrates in Mississippi. Hydrobiologia 573:155-172.

Bryce, S. A., G. A. Lomnicky, and P. R. Kaufmann. 2010. Protecting sediment-sensitive aquatic species in mountain streams through the application of biologically based streambed sediment criteria. Journal of the North American Benthological Society 29:657-672.

Carlisle, D. M., M. R. Meador, S. R. Moulton II, and P. M. Ruhl. 2007. Estimation and application of indicator values for common macroinvertebrate genera and families of the United States. Ecological Indicators 7:22-33.

Hawkins, C. P., J. R. Olson, and R. A. Hill. 2010. The reference condition: predicting benchmarks for ecological and water-quality assessments. Journal of the North American Benthological Society 29:312-343.

Horrigan, N., J. E. Dunlop, B. J. Kefford, and F. Zavahir. 2007. Acute toxicity largely reflects the salinity sensitivity of stream macroinvertebrates derived using field distributions. Marine and Freshwater Research 58:178-186.

Hubler, S., D. D. Huff, P. Edwards, and Y. Pan. 2016. The Biological Sediment Tolerance Index: Assessing fine sediments conditions in Oregon streams using macroinvertebrates. Ecological Indicators 67:132-145.

Huff, D. D., S. L. Hubler, Y. Pan, and D. L. Drake. 2008. Detecting shifts in macroinvertebrate assemblage requirements: implicating causes of impairment in streams. Oregon Department of Environmental Quality Watershed Assessment. Technical Report: DEQ06-LAB-0068-TR.

Lenat, D. R. 1993. A biotic index for the southeastern United States: Derivation and list of tolerance values, with criteria for assigning water-quality ratings. Journal of the North American Benthological Society 12:279-290.

Malmqvist, B. and S. Rundle. 2002. Threats to the running water ecosystems of the world. Environmental Conservation 29:134-153.

Meador, M. R., D. M. Carlisle, and J. F. Coles. 2008. Use of tolerance values to diagnose water-quality stressors to aquatic biota in New England streams. Ecological Indicators 8:718-728.

Norton, S. B., S. M. Cormier, M. Smith, and R. C. Jones. 2000. Can biological assessments discriminate among types of stress?: A case study from the eastern corn belt plains ecoregion. Environmental Toxicology and Chemistry 19:1113-1119.

Relyea, C. D., G. W. Minshall, and R. J. Danehy. 2012. Development and validation of an aquatic fine sediment biotic index. Environmental Management 49:242-252.

Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson, and R. H. Norris. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications 16:1267-1276.

Ter Braak, C. J. F, and L. G. Barendregt. 1986. Weighted averaging of species indicator values: its efficiency in environmental calibration. Mathematical Biosciences 78:57-72.

Tripole, S., E. A. Vallania, and M. del Carmen Corigliano. 2008. Benthic macroinvertebrate tolerance to water acidity in the Grande River sub-basin (San Luis, Argentina). Limnetica 27:29-38.

Turley, M. D., G. S. Bilotta, C. A. Extence, and R. E. Brazier. 2014. Evaluation of a fine sediment biomonitoring tool across a wide range of temperate rivers and streams. Freshwater Biology 59:2268-2277.

U.S. Environmental Protection Agency. 1994. Watershed protection: TMDL Note #2: Bioassessment and TMDLs. EPA/841/K-94/005a. Office of Water, U.S. Environmental Protection Agency, Washington, DC.

U.S. Environmental Protection Agency. 2006. Wadeable streams assessment: a collaborative survey of the nation's streams. EPA/641/B-06/002. Office of Water, U.S. Environmental Protection Agency, Washington, DC.

Yuan, L. L. 2006. Estimation and application of macroinvertebrate tolerance values. EPA/600/P-04/116F. U.S. Environmental Protection Agency, Washington, DC.

TABLES AND FIGURES

Table 1.1. Hypothetical example showing how a comparison of the environmental preferences of expected taxa to those of observed taxa could be used to diagnose the causes of biological stream impairment. Mean E = mean value of the expected taxa, Mean O = mean value of the observed taxa, Temp = water temperature (˚C) preference, and % Fine Sed. = preference for percentage of fine sediment in the streambed.

Taxon	Expected Taxa	Observed Taxa	Temp.	% Fine Sed.
A	X		6	15
B	X		10	5
C	X	X	14	45
D	X	X	7	50
E	X	X	15	40
F		X	8	80
G		X	10	75
		Mean E:	10.4	31
		Mean O:	10.8	58

Table 1.2. Spearman's rank correlation coefficients comparing tolerance values for percent fine sediment for 23 taxa across five studies (Yuan 2006, Carlisle et al. 2007, Relyea 2007, Huff et al. 2008, Bryce et al. 2010). The * indicates the correlation is significant at a Bonferroni-adjusted alpha of 0.005.

	Bryce	Huff	Yuan	Carlisle
Huff	0.68*			
Yuan	0.39	0.30		
Carlisle	0.34	0.41	0.63*	
Relyea	0.52	0.40	0.24	0.29

CHAPTER 2

REFINING METHODS FOR QUANTIFYING MACROINVERTEBRATE ESTIMATES OF PREFERENCE FOR USE IN STREAM BIOASSESSMENTS

Introduction

Reliable, quantitative tools are needed to identify the causes of biological impairment in the Nation's aquatic ecosystems. Sixty-seven percent of stream length in the United States is in either fair or poor biological condition (USEPA 2006). Potential causes of biological impairment in streams include excess nutrients, fine sediments, salinity, loss of riparian vegetative cover, and toxic compounds (Zweig and Rabeni 2001, Courtney and Clements 2002, USEPA 2006). However, we do not yet have reliable tools to identify the specific causes of biological impairment on a site-by-site basis. One potential way to diagnose the causes of impairment is to compare the specific environmental preferences or tolerances of the taxa in the biological community expected at a site to the taxa in the observed community. For example, observing more warm-water taxa (or fewer cool-water taxa) than expected implies that a stream has been thermally altered. Aquatic macroinvertebrates are good candidates for conducting this type of diagnosis because there are many taxa, taxa have different environmental preferences, they are found in almost all streams, and they integrate the effects of stressors over time (USEPA 2006). Stream ecologists have recently begun to compile information on the environmental preferences and tolerances of different macroinvertebrate taxa to support such causal analyses. Researchers have variously

referred to these preference estimates as tolerance values (Yuan 2006), indicator values (Carlisle et al. 2007), and occurrence values (Relyea et al. 2012). They have also used various analytical methods to estimate these values, including weighted averaging (Brandt 2001, Yuan 2006, Carlisle et al. 2007, Huff et al. 2008, Bryce et al. 2010), percentiles of occurrence or abundance along environmental gradients (Yuan 2006, Larsen et al. 2009, Relyea et al. 2012), non-linear regression (Yuan 2006), and identification of the value along an environmental gradient with the highest taxon densities or most presences (Yuan 2006).

Applying this approach to causal assessment assumes that one can generate reliable and repeatable taxon-specific estimates of environmental preferences/tolerances. Most estimates have been derived from field survey data. However, inconsistencies in these estimates among different researchers cast doubt on their general accuracy and utility (Figure 2.1, Table 1.2). For example, Huff et al. (2008) estimated that the mayfly genus *Ameletus* preferred stream substrates composed of 6% fine sediment, whereas Relyea (2007) and Yuan (2006) estimated that *Ameletus* preferred substrates with 60% and 20% fine sediment, respectively. These inconsistencies in estimates are not limited to fine sediment. Yuan (2006) estimated a temperature preference of 11°C for the stonefly genus *Yoraperla*, whereas Huff et al. (2008) and Brandt (2001) estimated preferences of 14°C and 9°C, respectively. The inconsistencies among these estimates imply that many or all estimates are unreliable, and that their use would often lead to inaccurate causal assessments.

Published estimates of preference and tolerance may not agree with one another for several possible reasons. Two of these reasons are confounding among environmental variables in the data used to estimate preferences, and the effects of interactions among environmental factors on taxon presence and abundance. Confounding is typically a problem when two or more environmental variables co-vary among the observations used to estimate preferences (Suter and Cormier 2013), which results in biased estimates of the relationship between biota and one or both environmental attributes. For example, at the spatial grain of stream reaches, the average amount of fine sediment in the streambed can co-vary with average temperature, nutrient concentrations, or flow velocity. This potential co-variance can make it difficult to interpret estimates of preference for any single factor (Zweig and Rabeni 2001; Yuan 2006, 2007, 2010). However, confounding effects might be minimized if environmental variables can be measured at a smaller spatial grain, rather than averaged across the reach (Larsen et al. 2009). Interactive effects occur when a taxon's response to one environmental factor depends on the level of another factor, such as a taxon's velocity preference changing with food availability. Interactive effects of environmental factors may thus influence the interpretation of a taxon's response to environmental disturbance (Matthaei et al. 2010, Piliere et al. 2014).

Incomplete sampling of a taxon's niche space can also lead to biased estimates of preference/tolerance. For example, if temperatures at sampled streams ranged between 5 and 15°C, then estimates of preference must fall within that ten-degree range. If the thermal niche of a sampled taxon actually spanned 10 to 30°C, the sample data would

necessarily lead to an underestimate of thermal preference. If the entire niche space of a taxon is sampled, the estimate of preference is more likely to be accurate. Some data sets previously used to estimate preferences/tolerances come from studies that sampled large portions of the United States (e.g. Yuan 2006, Carlisle et al. 2007). These large-scale studies almost certainly sampled most of the available environmental space in U.S. streams and should therefore encompass much, if not all, of the niche space of many taxa. Other field data sets used to estimate preferences/tolerances, however, come from studies that may have sampled only a portion of available niche space, such as just the range of conditions that occur within an individual state (e.g. Huff et al. 2008). These studies are more likely to produce biased estimates for at least some taxa.

Another potential problem with current estimates of preference is that they are based on habitat use, i.e., where target organisms are observed or collected. These estimates generally do not account for habitat availability. Estimates of preference derived solely from habitat-use data will likely be affected by the specific environmental conditions available at the time of sampling and thus may not actually indicate preference. Assuming that individuals primarily distribute themselves in response to differences in habitat quality, one can infer that a taxon prefers particular environmental conditions if the proportion of individuals found in those conditions is greater than the proportional availability of those conditions. Estimates of preference that incorporate both habitat-use and availability data are more independent of the sampled environment and should more reliably represent preferences in locations outside of the sampled area (Bovee 1986).

My research objectives were to determine: (1) if sampling of small patches within reaches can eliminate confounding in estimates of environmental preferences/tolerances derived from survey data, (2) if interactions among environmental factors affect estimates of preference/tolerance, (3) how strongly incomplete sampling of niche space affects estimates of preference/tolerance, and (4) if different methods of generating estimates produce comparable inferences. By addressing these objectives, I should be able to determine if taxon- and stressor-specific preference/tolerance estimates can be used reliably as the basis for causal assessment of biological impairment in streams.

In the rest of this thesis, I use the term "tolerance values" (or "TVs") to collectively refer to measures of environmental optima (preference) or tolerance limits estimated by different methods. However, I use the terms "preference" and "tolerance" when referring specifically about what each method actually measures.

Methods

Research design and study area

I collected patch-level (i.e., ~0.1 m^2) data for macroinvertebrates across a wide variety of streams. Streams were selected to represent the range of temperatures and salinities that occur in the region. These two habitat variables can vary markedly across reaches but are generally homogeneous within reaches. Sampled patches were nested within each stream reach to assess within-reach variation in substrate size and flow velocity. This sampling regime was specifically designed to minimize confounding between substrate size, which varies at both patch and reach scales, and temperature and

salinity (measured as specific electrical conductance, hereafter referred to as conductivity), both of which mainly vary across, and not within, reaches.

I sampled 45 streams in Idaho, Oregon, Utah, and Nevada between May and September 2010 (Figure 2.2; Table 2.1). Streams ranged in size from <1 to approximately 15 m wetted width. In each stream reach, I selected five to eleven small (0.09 m^2) patches that spanned the range of substrate sizes and flow velocities present at each site, which resulted in a total of 332 sampled patches.

Sampling methods

At each sampling location, I sampled macroinvertebrates and measured water temperature (˚C), conductivity (µS/cm), water velocity (m/s), and substrate size composition (mm) from each 0.09 m^2 sample patch. Macroinvertebrates and substrates (top 5 to 10 cm of the streambed) were collected with a 500-µm-mesh Surber sampler. The entire sample was then transferred to a bucket, and macroinvertebrates were elutriated from the substrates. Heavier invertebrates, such as caddisflies and snails, were removed from the substrates by hand. Macroinvertebrates were preserved in the field in 95% ethanol.

While in the field, I washed each sample's substrates through a series of sieves (64, 32, 16, 8, and 4 mm). All particles smaller than 4 mm, or a well-mixed subsample of these particles, were retained and later sieved through 2-, 1-, and 0.5-mm sieves in the lab. Sieves with mesh finer than 0.5-mm were not used because substrates smaller than 0.5-mm may have gone through the mesh of the Surber sampler when the sample was collected. I used graduated cylinders to measure, by displacement, the volume of each

substrate size fraction. The volumes of the largest rocks were approximated by measuring each rock's x-, y-, and z-axes and assuming the rocks were rectangular (length x width x height) or pyramidal (⅓ length x width x height) in shape. Volumes of each of the nine substrate size categories (>64 mm, >32-64 mm, >16-32 mm, >8-16 mm, >4-8 mm, >2-4 mm, >1-2 mm, >0.5-1 mm, and <0.5 mm) were converted into percentages of the total sample. By weighting the median particle size in each substrate size category (i.e., >8-16 mm, >4-8 mm, etc.) by the corresponding percent of total volume, I calculated the median particle size (d50) for each sampled patch. I also calculated the percentage of fine sediment (<2 mm; PCLT2) present in each patch.

Temperatures (Temp) and conductivities (Cond) at each patch were measured with an Extech EC400 ExStik® meter. Because values varied little within a reach, I used reach averages for these two variables in my analyses. Temperatures were also standardized to a common day and time of day by fitting observations to a CART (classification and regression tree) model ("rpart" library; R 2.13.0, R Development Core Team, 2011) and then using the model to adjust values to a standard day and time. CART models sequentially partition the dataset to "maximize differences on a dependent variable" (Hair et al. 1998). The independent variables that best partitioned the temperature data were the number of hours past 5 a.m. and the day of the year. Temperatures were standardized because samples were taken over a span of four months and at all times of day. Stream reach temperatures can fluctuate substantially within and across days.

Water velocity (Vel) was estimated at each patch as the average of the reciprocals of the times it took the leading and trailing edges of fluorescein dye to travel one meter in the water column above the patch. Dyes injected into a stream act in the same manner as the water molecules (Kilpatrick and Wilson 1989), so this method should adequately estimate the mean water-column velocity.

In the lab, approximately 100 individual non-chironomid macroinvertebrates from each sample were sorted and identified to the lowest practical taxonomic level. Abundances of the taxa in each sample were then scaled to density per square meter. Sixty-three taxa were chosen for analyses: 15 species, 41 genera, 4 families, and 3 higher-level taxa (Table 2.2; Appendix Table A.1). These 63 taxa each occurred in twenty or more patches. Chironomids were not included in these analyses because of the cost of identifying them beyond subfamily.

Analyses to address confounding

If patch-level sampling minimized confounding between environmental variables, correlations among variables measured at the patch scale should be substantially lower than correlations based on reach-scale average conditions. I calculated Spearman's rank correlation coefficients between temperature, conductivity, velocity, median particle size, and percent fines data at both patch and reach scales to determine if patch sampling minimized confounding among variables. Spearman's rank correlation coefficients are appropriate for datasets such as mine, where environmental variables are not normally distributed and relationships between some environmental variables are non-linear. I then compared the correlation coefficients generated from patch data with coefficients

generated from reach-averaged environmental data. Statistical significance was assessed after Bonferroni correction for multiple comparisons.

Analyses to address interactions

To identify associations between taxon densities and each of the variables measured in the study (temperature, conductivity, velocity, median particle size, and percent fine sediment), I created Random Forest models ("randomForest" library; R 2.13.0, R Development Core Team, 2011). Bivariate partial dependence plots (Cutler et al. 2007) allowed me to visualize density responses to joint variation in different paired combinations of environmental variables ("bivarpartialPlot" function in the "randomForest" library, R 2.13.0, R Development Core Team, 2011). I created bivariate partial dependence plots for twelve taxa (indicated by * in Table 2.2) that represented the range of living strategies and taxonomic orders (three mayflies, two stoneflies, two caddis flies, two true flies, one beetle, one snail, and one mite) in the data, which should therefore represent the general importance of interactions among the 63 target taxa. In each plot, if the response pattern to one variable changed as values of the second variable changed, I concluded that the two factors likely interacted to influence density and thus factor-specific estimates of preference. I also tried to use generalized linear mixed models to objectively test for the effects of interactions among environmental variables on taxon abundances, but these models generally failed to converge on a solution that described the non-linearity of taxon responses to the environmental variables.

Analyses to examine effects of incomplete niche sampling

To examine the effects of incomplete niche sampling, I compared estimates across sub-regions of environmental space. The 332 patches were assigned to one of six bins based on two levels of temperature (<9.5°C and ≥ 9.5°C) and three levels of conductivity (<70 μS/cm, 70 to 225 μS/cm, and >225 μS/cm) (Figure 2.3). Each of these six bins thus represented a portion of the environmental space that existed across all of the sampled sites. Weighted-average estimates of preference for each of the 63 taxa for each temperature/conductivity bin were calculated as:

$$WA_j = \sum Y_{ij} x_i / \sum Y_{ij}, \quad (1)$$

where Y is the density (individuals/m^2) of taxon j, and x is the value of the environmental variable of interest at site i. For each environmental variable, weighted-average preference estimates for each taxon were plotted against the mean values for that variable calculated from the stream-reach data in each bin. If the weighted-average preference estimates varied proportionally with the mean values, I concluded that sampling only a portion of environmental space could bias estimates of preference.

Effects of analytical method on preference estimates

For each of the 63 taxa, I estimated preference or tolerance for the five different environmental variables in six different ways: (1) with abundance-weighted averages (WA), (2) as the position along an environmental gradient at which peak density was observed (DPeak), (3) as the position along an environmental gradient at which the most presences were observed (PPeak), (4) as the position along an environmental gradient

below which 75% of the individuals in a taxon are present (75th), (5) with habitat-selectivity indices (SI), and (6) with ratios of the number of observed individuals to expected individuals (O:E). Methods 1, 2, 3, 5, and 6 estimate taxa preference, whereas method 4 estimates a taxon's tolerance limit.

The first four estimation methods were based on the presence or density of taxa. I used equation 1 to estimate weighted averages for all five variables based on data from all 332 patches. Density-environment relationships (method 2) were created by plotting the densities of each taxon against values of each of the environmental factors, and preference was inferred as the point along the environmental gradient where the taxon's estimated density was highest. If densities were equally high at two points along the gradient, I estimated preference as the value equidistant between the two points. I also created dot-density plots that showed the number of times at least one individual in a taxon was observed at intervals along each gradient (method 3). Preference was inferred as the median of that interval along the gradient where the most presences occurred. I also estimated the 75^{th} percentile of occurrence along each of the five environmental gradients (method 4). Lenat (1993) and Relyea et al. (2012) found 75% to be the most effective threshold to describe a taxon's tolerance to an environmental variable by creating the greatest separation between tolerant and intolerant species.

I also calculated two preference indices based on the use and availability of specific habitat conditions across all sampled patches. In method 5, I used the relativized electivity index (E*), which incorporates both habitat use (taxon presence) and habitat

availability to estimate habitat preference (Vanderploeg and Scavia 1979). E* is calculated as:

$$E^* = [(R/\sum R)-(1/n)]/[(R/\sum R)+(1/n)], \quad (2)$$

where *R* is the ratio of proportions of used to available habitat in each class interval, and *n* is the number of class intervals (Gibbins et al. 2002). This method requires that the environmental gradient be divided into bins (Table 2.3). E* was then calculated for each bin. For these analyses, I divided the temperature, conductivity, velocity, and median particle-size gradients into five equal-interval bins. However, for the percent fines gradient, equal-interval binning resulted in extremely unequal numbers of samples in each bin. I therefore divided the percent-fines gradient into five bins whose intervals progressively doubled in size. These divisions seemed to be biologically meaningful and mirrored methods used in other studies (e.g. Gibbins et al. 2002, Sugiyama and Goto 2002). E* can range from -1 to 1. A value of 0 indicates that a taxon selects a habitat at random, whereas values >0 and <0 indicate preference or avoidance, respectively. Once the most-preferred bin was identified, I estimated preference (SI) as the median environmental value of that bin.

The other method incorporating both habitat use and availability (method 6) compares the observed number of individuals to the number expected in each of *n* habitat categories, assuming individuals were distributed among habitat categories in proportion to their availability. The expected number of individuals in each habitat category was estimated as the total number of individuals in a taxon multiplied by the proportional

availability of each habitat category. The habitat categories were the same bins used in the calculation of E* (Table 2.3). I then used chi-square tests to determine if the observed number of individuals was significantly different from the expected number in each bin (Storaas and Wegge 1987). The magnitude (and therefore the likelihood of significance) of chi-square values increases as the magnitude of values being compared increases. To avoid inflating significance tests, I therefore used sample counts rather than extrapolated density estimates. I then calculated observed to expected (O:E) ratios, and inferred that the bin with the greatest O:E value was most preferred. The median value of that bin was used as the taxon's preference.

Comparing estimates of preference

To quantify the degree to which tolerance values based on the six methods agreed, I calculated Spearman's rank correlation coefficients across the 63 taxa for each of the five habitat variables. Statistical significance was assessed based on Bonferroni corrections for multiple comparisons. In addition, I used non-metric multi-dimensional scaling (NMDS; Primer statistical package, v. 6, 2006) to produce a two-dimensional "map" of the proximity of the 30 tolerance value combinations (five factors x six methods) to one another in tolerance value space (Hair et al. 1998). Each tolerance value pair (e.g., WA-Temp for taxon A) was standardized (normalized) using the mean and standard deviation of the distribution. I used Euclidean distance as the measure of similarity between all pairs of tolerance value types. The position of the different method-type combinations in ordination space should reveal if different types of tolerance values

are discriminating among different stressor signals and how strongly method affects, or confounds, those signals.

Results

Ranges of environmental conditions

Patch environments varied markedly among sites (Table 2.4). For example, date (July 15)- and time (11:30)-adjusted temperature varied from 6 to 18°C, conductivities ranged from 24 to 542 μS/cm, water velocity varied from 0 to 1.6 m/s, median particle sizes ranged from 1 to 88 mm, and percent fine sediment ranged from 0 to 97% (Figure 2.4).

Assessment of confounding

Reach-level confounding was not always as strong as expected. There was not a significant correlation between temperature and fine sediment tolerance values in my data (Table 2.5). However, two pairs of reach-scale environmental variables were significantly correlated: percent fines and median particle size (-0.76), and temperature and conductivity (0.43). Correlations based on patch-scale data were lower than reach-scale correlations in five of ten cases. The absolute values of four other patch-level correlation coefficients were higher than the reach-level correlations, but not substantially so (Table 2.5). More patch-level correlations were statistically significant as expected because of a larger sample size.

Assessment of interactions

I found no evidence that interactions between factors affected the abundance of individual taxa. Visual interpretation of all 120 bivariate partial-dependence plots, which show the response of each of the twelve target taxa to all possible pairwise combinations of the five environmental variables, indicated that the pattern of response to one of the variables was consistent across all values of the second variable. For example, the response of *Ameletus* to temperature exhibited no obvious change with variation in percent fines (Figure 2.5). Therefore, the effects of temperature and percent fines on *Ameletus* abundance are additive, rather than interactive (Cutler et al. 2007). Other bivariate partial dependence plots (Figures 2.6 through 2.9) indicated similar additive effects of environmental variables on *Simulium*, *Drunella doddsii*, *Physa*, and *Optioservus quadrimaculatus* abundances.

Assessment of incomplete sampling of niche space

Weighted-average estimates of habitat preference varied substantially across temperature/conductivity bins. For example, conductivity preferences based on weighted-averages for the mayflies *Epeorus* and *Acentrella* ranged ten-fold, from 30 to 300 μS/cm and 30 to 375 μS/cm, respectively. Mean conductivities across the bins used in analyses ranged eight-fold (50 to 400 μS/cm), and preference estimates were strongly related to the mean conductivities within bins (Figures 2.10 and 2.11). The relationships between mean conductivities in the temperature/conductivity bins and preference estimates for temperature and conductivity were strong for all taxa (Figures 2.12a and 2.12b).

The relationships between weighted-average preference estimates and mean values of velocity, median particle size, and percent fines in each temperature/conductivity bin were less strong and inconsistent across taxa (Figure 2.13). For example, the percent fines weighted-average preference estimates for the mayfly *Serratella* ranged from 5 to 16%, and from 4 to 17% for the beetle *Optioservus*. Mean percent fines across the bins ranged two-fold (7 to 13%). A strong linearly increasing relationship existed between percent fines preference estimates and mean percent fines within the bins for *Serratella* (Figure 2.14), but no relationship existed between the two variables for *Optioservus* (Figure 2.15). Bins were based on ranges of temperature and conductivity, and the other environmental variables ranged widely and overlapped extensively within each bin. Therefore, preference estimates for these other environmental variables did not consistently increase as bin averages increased. Estimates for some taxa were similar across bins: *Turbellaria* and *Diphetor hageni* (median particle size), *Zapada* (velocity), *Dicranota* and *Epeorus* (percent fine sediments). Estimates for other taxa decreased as bin averages increased: *Dicranota* and *Paraleptophlebia* (median particle size), *Antocha monticola*, *Glossosoma*, and *Hydroptilidae* (velocity).

Assessment of different methods of estimating preference

The consistency, or lack thereof, among the tolerance values generated across methods depended upon the taxon and the environmental variable being considered (Appendix Tables A.2 through A.6). For example, the variation in temperature TV estimates for the mayfly *Centroptilum* was relatively small (12 to 13.8°C), whereas the

variation among estimates for the dipteran *Limnophila* ranged from 6.3 to 14°C. Similarly, variation in conductivity TVs for the stonefly *Zapada* ranged from 25 to 400 µS/cm, whereas the conductivity TVs for the snail *Physa* were much less variable, ranging from 374 to 423 µS/cm. Because the estimate generated from the 75th percentile of occurrence is actually a measure of tolerance rather than preference, one would expect these estimates to be consistently greater than estimates generated with the other five methods. For some taxa, the "75th" estimate was the greatest of the six estimates (e.g., temperature estimates for *Dicosmoecus* ranged from 6.3°C (SI) to 9.3°C (75th) and particle size estimates for *Limnophila* ranged from 6.7 mm (O:E and SI) to 35.4 mm (75th). However, this was not true for all taxa. For example, the "75th" percent fines estimate for *Dicranota* was that taxon's second lowest estimate, and the "75th" conductivity estimate for *Arctopsyche grandis* was the second highest estimate. For particle size estimates, the 75th percentile of occurrence can be misleading. If these estimates are to be interpreted as a measure of tolerance to fine sediments (the stressor), a 25th percentile of occurrence may be more intuitive. However, I chose not to convert these estimates from 75th to 25th percentiles so that direct comparisons of the six types of estimates could be made.

Between-method temperature and conductivity estimates were more consistently correlated than estimates for velocity, particle size, and percent fine sediment (Table 2.6, Appendix Figures A.1 through A.5). Thirteen of fifteen between-method Spearman's rank correlation coefficients were significant for temperature, and twelve were significant for conductivity, though some pairs were weakly correlated (e.g., 0.37 between 75th

percentiles and the peak of the density curve for temperature). Significant correlations also existed between 8, 10, and 8 pairs of methods for velocity, median particle size, and percent fines, respectively. Correlations between estimates generated from weighted averaging and O:E ratios were consistently the highest across the five environmental variables (0.85, 0.82, 0.81, 0.78, and 0.85 for temperature, conductivity, velocity, median particle size, and percent fines, respectively). In contrast, correlation coefficients between estimates based on the peak of the presence curve and any other method were consistently the lowest across the five variables (0.32, 0.29, 0.03, 0.21, and 0.00 for temperature, conductivity, velocity, median particle size, and percent fines, respectively).

NMDS Ordination

The NMDS ordination (stress = 0.14, two dimensions) separated tolerance values based on habitat attributes (Figure 2.16). However, methods for estimating TVs for temperature and conductivity grouped next to each other in ordination space, as did methods for estimating TVs for median particle size and velocity. Methods for estimating TVs for temperature and median particle size formed relatively tight clusters, whereas the remaining clusters were larger and influenced by apparent outliers. In particular, PPkVel and PPkPCLT2 were somewhat isolated from the rest of the methods used to estimate TVs for velocity and percent fine sediment, respectively.

Discussion

Reliable, quantitative tools are needed to diagnose the causes of biological impairment in streams. Comparing stressor-specific tolerance values of the

macroinvertebrate assemblage expected at a potentially degraded site to the assemblage observed at that site may provide insight into the causes of degradation. However, researchers and managers must be able to produce repeatable estimates of stressor-specific preference or tolerance values for each taxon before they can be used reliably at multiple locations and spatial scales. Unfortunately, consistency in estimates derived from survey data has not yet been achieved (Figure 2.1). Reasons for current inconsistencies in these estimates could include confounding, interactions, incomplete sampling, and the method used to estimate preference.

Confounding

The issue of confounding makes it very difficult to determine exactly what tolerance values measure. For example, do temperature TVs actually describe a taxon's preference/tolerance for temperature? Or do they describe a taxon's preference/tolerance for some stressor that co-varies with temperature? This uncertainty affects the diagnostic power of tolerance values and makes them difficult to interpret. This field study, in which small patches of stream reaches were sampled, was designed to minimize confounding from TVs. However, statistically significant correlations among habitat attributes at both the reach and patch scales suggest that sampling at the patch scale may not eliminate confounding (though many correlation coefficients were relatively small; Table 2.5). Two substrate variables, median particle size and percent fine sediment, were highly correlated and, therefore, likely to be highly confounded, whereas temperature and conductivity were less so. The high negative correlation between median particle size and percent fines was expected because greater quantities of fine sediment in a sample

will decrease that sample's median particle size. Contrary to the results of this study, Larsen et al. (2009) found that reducing the scale of sampling from reach- to patch-scale ($0.16m^2$) removed much of the confounding among those environmental variables that were altered by land use, such as hydrology, nutrient release, and energy flux. However, Larsen et al. also reported that some habitat attributes were significantly correlated at the patch scale, including sediment cover and water depth. Because Larsen et al. did not include temperature or conductivity in their measured habitat attributes, it is difficult to directly compare my results with theirs.

The results of this study reinforce concerns expressed by other authors (Zweig and Rabeni 2001, Bressler et al. 2006, Yuan 2006, Carlisle et al. 2007, Larsen et al. 2009, Suter and Cormier 2013) that confounding of environmental variables can make it difficult to isolate the responses of taxa to individual variables when generating tolerance values. These authors either simply acknowledged that confounding is a potential weakness of their estimates (Carlisle et al. 2007), or they used various strategies to minimize or circumvent confounding, including statistically controlling for sources of confounding (Yuan 2006), sampling streams that varied only in one factor (Zweig and Rabeni 2001), or sampling macroinvertebrates and habitat variables at smaller-than-reach scales (Larsen et al. 2009). Another alternative is to develop tolerance values based on a generalized stressor gradient that includes a variety of chemical and physical stressors (e.g., Bressler et al. 2006), but this approach would exclude the use of these values to diagnose specific causes of impairment.

The NMDS ordination also revealed some degree of confounding in tolerance values (Figure 2.16). Temperature and conductivity TVs based on different methods clustered together in ordination space suggesting that taxon responses to these two variables cannot be easily isolated with field data. The same was true for median particle size and velocity. However, contrary to expectation, TVs for percent fine sediment were distinctly segregated from those for median particle size in ordination space, despite the strong correlation between these two factors both within and across streams. This observation suggests that correlated variables should not necessarily be eliminated from an analysis without first examining the relationships between TVs for the two habitat attributes. Taxon preferences for the two correlated environmental variables may be independent from one another and both may be relevant in a bioassessment.

Eliminating confounding from field survey data is difficult. The natural variability of conditions in streams, as well as anthropogenic impacts, make teasing apart macroinvertebrate responses to one stressor nearly impossible. However, experiments may corroborate (or invalidate) estimates of preference/tolerance from field surveys. In experiments, researchers can control for the effects of potentially confounding factors, and TVs derived from experimental work should accurately describe a taxon's preference or tolerance to a single stressor. Experimental manipulations, however, are resource-intensive, and the act of removing and transporting organisms from their natural habitat to an experimental setup may introduce other unknown, confounding effects (experimental stress, etc.). However, if experiments *do* corroborate TVs derived from

field survey data, researchers and managers can be more confident in the interpretation of field-derived TVs.

Interactions

Interactions among environmental variables imply that the interpretation of a given tolerance value is dependent on the other environmental conditions that occur at each assessed site, which would greatly compromise the general applicability of taxon- and stressor-specific TVs. However, no evidence of interactions among environmental variables were found in this study (Figures 2.5-2.9), and interactions therefore may not be as problematic as confounding is. These results contrast with those of Piliere et al. (2014), who claim interactions among environmental factors influenced their ability to interpret relationships between several macroinvertebrate metrics and different predictor variables. However, many significant interactions in their study were found between surrogates for more biologically meaningful factors. For example, the interaction they report between altitude and latitude was interpreted as an interaction between water temperature (or substrate type) and large-scale biogeographical patterns (such as precipitation), respectively. Interactions between these non-specific environmental factors are difficult to interpret because macroinvertebrates could be responding to any one of several factors that co-vary with geographic gradients.

Other authors (e.g., Townsend et al. 2008, Matthaei et al. 2010) report evidence of interactions among variables such as current velocity, sediment, and nutrients influencing macroinvertebrate richness and abundance. However, these interactions were more pronounced in their experimental streams than in their field survey data. These results, in

conjunction with the lack of evidence for interactions in my analyses, suggest that experimental manipulations may be capable of detecting potential interactions, but that these interactions may be either less important or obscured in field data that are simultaneously influenced by multiple variables (Townsend et al. 2008, Downes 2010, Piliere et al. 2014).

Incomplete sampling of niche space

If tolerance values are estimated from field survey data that only covers a portion of a taxon's niche space, then those TVs will likely be inaccurate, depending on how much and what part of the environmental gradient is sampled. My results clearly show that sampling only a portion of a taxon's niche space can severely bias TV estimates. This potential for bias was obvious for all taxa in subsets of environmental space (Figures 2.12a, 2.12b, 2.13). This bias was most notable for temperature and conductivity because I intentionally parsed environmental space based on these two factors. However, bias was also detected in the relationships between TVs and velocity, median particle size, and percent fines, though these relationships were not consistent across taxa. These results indicate that reliable estimates of a taxon's environmental preferences/tolerances can be obtained from field data only when the sampled environment includes that taxon's entire niche – a conclusion also reached by Zettler et al. (2013) in their analysis of salinity preferences of six species of marine invertebrates.

In an effort to decrease sampling bias, I sampled as many streams across as wide a range of environmental conditions as possible. The ranges of temperature and percent fine sediment measured in this study were similar to those in other regional-scale studies

(comprising one or a few states) in the western United States (Table 2.4; Brandt 2001, Huff et al. 2008, Relyea et al. 2012). However, the ranges of the five variables in this study may not encompass the entire niche of every taxon sampled. Plots of taxon densities versus environmental gradients suggest that the niches of some sampled taxa include colder temperatures, higher conductivities, and higher velocities than were sampled. However, this was true for only a few taxa. Also, the distributions of the five variables were not uniform (Figure 2.4), which could undermine the accuracy of the estimates (Yuan 2006). For example, when using weighted averaging, an unequal distribution of samples across an environmental gradient will result in the unequal weighting of each location along that gradient, biasing the estimate away from the true value. Tolerance values derived from nationwide studies, such as the U.S. Geological Survey's National Water-Quality Assessment Program (Carlisle et al. 2007) and the Environmental Protection Agency's Environmental Monitoring and Assessment Program (Yuan 2006, Bryce et al. 2010), are more likely to encompass the entire niche space of the sampled taxa because the percent of taxa whose niches are fully represented likely increases as the geographic scope of the study area increases.

The problem of incomplete niche sampling can be addressed in two ways: (1) sample a very wide range of field conditions and plot macroinvertebrate abundances along each environmental gradient to ensure that the entire niche of each taxon was sampled, or (2) conduct controlled experiments for each stressor and include values of that stressor at the extremes of what the taxon experiences in the field. Researchers and

managers can then be more confident that TV estimates represent meaningful ecological responses.

Methods of estimating preference

The use of different methods to estimate tolerance values could affect the comparability of the estimates (because different methods scale estimates in different ways, are more precise or accurate than others, etc.). The six methods of estimating macroinvertebrate habitat preference or tolerance presented in this study lacked numerical consistency and precision (Appendix Tables A.2 through A.6). However, there is evidence that, despite the inconsistencies, the inferences that result from using any of these methods may be similar.

Although correlation coefficients between TVs calculated by the different methods ranged widely (-0.11 to 0.85; Table 2.6), the NMDS ordination suggested that consistent assemblage-wide signals emerged regardless of the method used. The fact that estimates produced for different variables grouped together in ordination space, rather than by method, suggests that the method used to estimate TVs is not the major source of variation in the comparisons I conducted (Figure 2.16). The two more-dispersed clouds of estimates in the ordination would have been more compact if the estimates based on the peak of the presence curve were removed. This observation suggests that PPeak estimates may be fundamentally different from the other five types of estimates (see below).

Weighted averaging, which defines a taxon's ecological optimum as the central value of that taxon's distribution, is one of the most popular and simplest methods used to

generate preference estimates (ter Braak and Barendregt 1986, Huff et al. 2008, Cristobal et al. 2014). However, a weighted average does not necessarily correspond to the peak taxon density if the taxon-environment relationship is not symmetrical and unimodal (Yuan 2006, Anderson 2008, Cristobal et al. 2014). Weighted averages can also be heavily influenced by outliers, which can either under- or overestimate estimates of preference (Cristobal et al. 2014). Despite the known limitations of the weighted averaging method, weighted average preference estimates were significantly correlated with estimates generated from other methods (Table 2.6). As inferred from the NMDS ordination, these significant correlations also suggest that relative differences in preference values for different factors are somewhat insensitive to method of estimating preference. However, high correlation coefficients between methods do not necessarily indicate that these different methods produce accurate estimates of a taxon's preference or tolerance for a given environmental factor. High correlations between methods could also simply reflect that all methods may be sensitive to the same issues of confounding, interaction, or incomplete sampling.

Preference estimates based on presence curves (PPeak) were not highly correlated (and rarely significantly correlated) with the other five methods (Table 2.6). PPeak estimates were derived by visually assessing the relationship between an environmental gradient and the number of instances a macroinvertebrate taxon was present in a sample. PPeak estimates can be heavily influenced by a single sample site, which can skew estimates of preference away from the "true" preference. This issue may be especially problematic for estimates derived from studies with few observations. Also, estimates

based on presence data do not take into account the relative abundance of taxa (Stark 1998). One individual has the same influence on an estimate of preference as 1000 individuals, and this equal influence may result in skewed estimates. For example, in this study, *Ephemerella inermis/dorothea* was assigned a PPeak conductivity estimate of 400 µS/cm. However, the greatest density of this species occurred at higher conductivities, and a conductivity TV estimate based on density would be closer to 500 µS/cm. A value of 500 µS/cm is more consistent with the other estimates generated from density data (507, 500, 423, and 481 µS/cm for WA, DPeak, O:E, and 75th estimates, respectively). A few observations of *Emphemerella inermis/dorothea* at low conductivity levels skewed the PPeak estimate.

Preference estimates based on density/environment curves (DPeak) can also be influenced by outliers. For example, the position along the percent fine sediment gradient at which peak Oligochaeta density was observed was at 95%. However, if that one sampling site was removed from the analysis, peak Oligochaeta density was estimated as 10% fine sediment. A value of 10% is more consistent with estimates for Oligochaeta based on the other five methods (34.1, 37.5, 37.5, 2.0, and 13.6 percent for WA, O:E, SI, PPeak, and 75th, respectively).

Estimates generated with the habitat selectivity index (SI) method did not make much ecological sense (Appendix Tables A.2 through A.6). For example, the SI method assigned 36 of the 63 taxa in this study to a d50 preference estimate of 75.6 mm. Of the remaining 27 taxa, 26 were assigned a d50 estimate of 6.7 mm (see Appendix Table A.5). It is ecologically unlikely that so many diverse taxa share the same preference for particle

size. The SI method was based on taxon presences, so SI estimates likely suffer from the same limitations as PPeak estimates. On the other hand, O:E estimates were based on taxon densities, which may result in more accurate estimates of preference. The SI and O:E methods are similar in that they are less likely to be skewed by the relative availability of specific environmental conditions found at sampled streams and are more likely to reflect habitat choices made by taxa (Bovee 1986, Heggenes 1990). Both methods require a binning process, which forces preference estimates to be one of *n* numbers, where *n* is the number of bins. For example, a macroinvertebrate with an actual conductivity preference of 360 µS/cm would likely be assigned an SI or O:E estimate of either 295 or 423 µS/cm, because those are the median values of the two bins closest to the "true" preference. However, using the median value of the bin as the estimate of preference may lessen the influence of outliers within each bin.

The final estimation method used in this study was the 75th percentile of occurrence. These estimates are not meant to reflect a macroinvertebrate's preference for an environmental variable, but rather its tolerance (i.e., the point along the environmental gradient where the macroinvertebrate population begins to decline (Yuan 2006)). In this study, preference estimates based on the 75th percentile of occurrence were significantly correlated with most other estimation methods across all habitat variables (Table 2.6), indicating that these estimates of tolerance would likely produce the same environmental signal in a causal analysis.

In summary, tolerance values estimated using the six different methods were numerically inexact, but these methods may produce comparable inferences in a causal

analysis. However, I suggest that some methods are better than others. For instance, I do not recommend the PPeak or SI methods for estimating TVs. These two methods made the least ecological sense and are heavily influenced by sites where few individuals are present. In contrast, the DPeak estimates can be heavily influenced by single sites with large numbers of individuals. Of the six methods used in this study, I recommend WA, O:E, or 75th. These three methods were most highly correlated with one another, suffered from the fewest biases, and made sense ecologically.

Conclusions

Stressor-specific estimates of macroinvertebrate tolerance values have the potential to be invaluable tools to diagnose the causes of biological impairment in streams, but their performance needs to be verified before researchers and managers can have much confidence in them. These estimates should reflect the true environmental preference/tolerance of the taxon for which they are derived, but there is no standard that can be used to judge the accuracy of field-derived TVs. At minimum, stressor-specific estimates derived from different studies should be consistent. Unfortunately, a consensus across studies does not yet exist. The results of this study suggest that incomplete sampling of a taxon's niche space may have the most profound and obvious effect on tolerance values, whereas interactions among environmental variables may have the least effect. Significant correlations between environmental variables suggest that confounding likely has some effect on TVs, but the magnitude of the effect is unclear. Also, though each method used to estimate tolerance values has limitations, TVs derived

from different methods may reveal the same environmental signal, and there may be little reason to prefer one method over another in a causal analysis. The set of tolerance values that are currently available need to be validated before researchers and managers rely on them as tools for causal assessment. Experimental studies should help validate field-derived estimates of preference/tolerance and improve our confidence in them. To further increase our confidence in estimates, researchers should continue to generate TVs for as many different taxa and stressors as possible.

REFERENCES

Anderson, M. J. 2008. Animal-sediment relationships re-visited: Characterising species' distributions along an environmental gradient using canonical analysis and quantile regression splines. Journal of Experimental Marine Biology and Ecology 366:16-27.

Bovee, K. D. 1986. Development and evaluation of habitat suitability criteria for use in the Instream Flow Incremental Methodology. Instream Flow Information Paper 21. U.S. Fish and Wildlife Service Biological Report 86(7). 235 pages.

Bressler, D. W., J. B. Stribling, M. J. Paul, and M. B. Hicks. 2006. Stressor tolerance values for benthic macroinvertebrates in Mississippi. Hydrobiologia 573:155-172.

Brandt, D. 2001. Temperature preferences and tolerances for 137 common Idaho macroinvertebrate taxa. Idaho Department of Environmental Quality. Coeur d'Alene, Idaho.

Bryce, S. A., G. A. Lomnicky, and P. R. Kaufmann. 2010. Protecting sediment-sensitive aquatic species in mountain streams through the application of biologically based streambed sediment criteria. Journal of the North American Benthological Society 29:657-672.

Carlisle, D. M., M. R. Meador, S. R. Moulton II, and P. M. Ruhl. 2007. Estimation and application of indicator values for common macroinvertebrate genera and families of the United States. Ecological Indicators 7:22-33.

Courtney, L. A. and W. H. Clements. 2002. Assessing the influence of water and substratum quality on benthic macroinvertebrate communities in a metal-polluted stream: an experimental approach. Freshwater Biology 47:1766-1778.

Cristobal, E., S. V. Ayuso, A. Justel, and M. Toro. 2014. Robust optima and tolerance ranges of biological indicators: a new method to identify sentinels of global warming. Ecological Research 29:55-68.

Cutler, D. R., T. C. Edwards, Jr., K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler. 2007. Random forests for classification in ecology. Ecology 88:2783-2792.

Downes, B. J. 2010. Back to the future: little-used tools and principles of scientific inference can help disentangle effects of multiple stressors on freshwater ecosystems. Freshwater Biology 55(Suppl. 1):60-79.

Gibbins, C. N., H. J. Moir, J. H. Webb, and C. Soulsby. 2002. Assessing discharge use by spawning Atlantic salmon: A comparison of discharge electivity indices and PHABSIM simulations. River Research and Applications 18:383-395.

Hair, J. F., R. E. Anderson, R. L. Tatham, and W. C. Black. 1998. *Multivariate Data Analysis*. 5th edition. Upper Saddle River, New Jersey: Prentice Hall.

Heggenes, J. 1990. Habitat utilization and preferences in juvenile Atlantic salmon (Salmo salar) in streams. Regulated Rivers: Research and Management 5:341-354.

Huff, D. D., S. L. Hubler, Y. Pan, and D. L. Drake. 2008. Detecting shifts in macroinvertebrate assemblage requirements: implicating causes of impairment in

streams. Oregon Department of Environmental Quality Watershed Assessment. Technical Report: DEQ06-LAB-0068-TR.

Kilpatrick, F. A. and J. F. Wilson, Jr. 1989. Chapter A9: Measurement of time of travel in streams by dye tracing. Techniques of Water-Resources Investigations of the United States Geological Survey, Book 3.

Larsen, S., I. P. Vaughan, and S. J. Ormerod. 2009. Scale-dependent effects of fine sediments on temperate headwater invertebrates. Freshwater Biology 54:203-219.

Lenat, D. R. 1993. A biotic index for the southeastern United States: Derivation and list of tolerance values, with criteria for assigning water-quality ratings. Journal of the North American Benthological Society 12:279-290.

Matthaei, C. D., J. J. Piggott, and C. R. Townsend. 2010. Multiple stressors in agricultural streams: interactions among sediment addition, nutrient enrichment, and water abstraction. Journal of Applied Ecology 47:639-649.

Piliere, A., A. M. Schipper, T. M. Breure, L. Posthuma, D. de Zwart, S. D. Dyer, and M. A. J. Huijbregts. 2014. Unraveling the relationships between freshwater invertebrate assemblages and interacting environmental factors. Freshwater Science 33:1148-1158.

Relyea, C. D. 2007. Fine inorganic sediment effects on stream macroinvertebrates. Ph.D. Dissertation. Idaho State University, Pocatello, Idaho.

Relyea, C. D., G. W. Minshall, and R. J. Danehy. 2012. Development and validation of an aquatic fine sediment biotic index. Environmental Management 49:242-252.

Stark, J. D. 1998. SQMCI: a biotic index for freshwater macroinvertebrate coded-abundance data. New Zealand Journal of Marine and Freshwater Research 32:55-66.

Storaas, T. and P. Wegge. 1987. Nesting habitats and nest predation in sympatric populations of capercaillie and black grouse. The Journal of Wildlife Management 51:167-172.

Sugiyama, H. and A. Goto. 2002. Habitat selection by larvae of a fluvial lamprey, Lethenteron reissneri, in a small stream and an experimental aquarium. Ichthyological Research 49:62-68.

Suter II, G. W. and S. M. Cormier. 2013. A method for assessing the potential for confounding applied to ionic strength in Central Appalachian streams. Environmental Toxicology and Chemistry 32:288-295.

Ter Braak, C. J. F, and L. G. Barendregt. 1986. Weighted averaging of species indicator values: its efficiency in environmental calibration. Mathematical Biosciences 78:57-72.

Townsend, C. R., S. S. Uhlmann, and C. D. Matthaei. 2008. Individual and combined responses of stream ecosystems to multiple stressors. Journal of Applied Ecology 45:1810-1819.

Vanderploeg, H. A. and D. Scavia. 1979. Calculation and use of selectivity coefficients of feeding: zooplankton grazing. Ecological Modeling 7:135-149.

U.S. Environmental Protection Agency. 2006. Wadeable streams assessment: a collaborative survey of the nation's streams. EPA/641/B-06/002. Office of Water, U.S. Environmental Protection Agency, Washington, DC.

Yuan, L. L. 2006. Estimation and application of macroinvertebrate tolerance values. EPA/600/P-04/116F. U.S. Environmental Protection Agency, Washington, DC.

Yuan, L. L. 2007. Using biological assemblage composition to infer the values of covarying environmental factors. Freshwater Biology 52:1159-1175.

Yuan, L. L. 2010. Estimating the effects of excess nutrients on stream invertebrates from observational data. Ecological Applications 20:110-125.

Zettler, M. L., C. E. Proffitt, A. Darr, S. Degraer, L. Devriese, C. Greathead, J. Kotta, P. Magni, G. Martin, H. Reiss, J. Speybroeck, D. Tagliapietra, G. Van Hoey, and T. Ysebaert. 2013. On the myths of indicator species: Issues and further consideration in the use of static concepts for ecological applications. PLOS ONE 8: e78219.

Zweig, L. D. and C. F. Rabeni. 2001. Biomonitoring for deposited sediment using benthic invertebrates: a test on 4 Missouri streams. Journal of the North American Benthological Society 20:643-657.

TABLES AND FIGURES

Table 2.1. Locations, dates, reach-scale environmental data, and numbers of taxa found in each of the 45 sampled streams. OR = Oregon, ID = Idaho, NV = Nevada, UT = Utah.

Stream	Latitude	Longitude	County	State	Date Sampled (in year 2010)	Number of Patches Sampled	Estimated July Temp (°C)	Conductivity (µS/cm)	Number of Taxa Found (out of the 63 taxa analyzed)
Vance	44.28476	-118.96852	Grant	OR	May 27	5	8.6	432	31
East Fork Beech	44.51119	-118.96860	Grant	OR	May 28	8	8.5	108	21
Fields	44.37769	-119.31462	Grant	OR	May 29	8	8.5	237	39
Middle Fork Canyon	44.21265	-118.84288	Grant	OR	May 30	8	9.2	63	36
Murderers	44.26348	-119.28441	Grant	OR	May 31	6	12.1	486	28
South Heglar Canyon	42.40720	-113.06970	Cassia	ID	June 4	8	13.0	423	17
Sublett	42.32587	-113.01241	Cassia	ID	June 5	8	13.8	542	18
Dorsey	42.05012	-115.52127	Owyhee	ID	June 10	8	12.1	184	12
Lakefork	42.33875	-113.04479	Cassia	ID	June 13	8	12.6	481	24
South	43.91493	-113.03259	Butte	ID	June 22 & 23	8	10.4	394	31
Badger	44.08934	-113.16721	Butte	ID	June 23 & 24	6	8.9	272	20
Squaw	44.12846	-113.38943	Butte	ID	June 25	7	9.7	324	29
Rooks	43.65122	-114.51348	Blaine	ID	June 26 & 27	8	8.7	221	28
Red Warrior	43.63840	-114.49580	Blaine	ID	June 27 & 28	7	7.9	107	30
Sand	41.85349	-115.76110	Elko	NV	July 7 and 8	8	10.9	58	19
Columbet	42.00974	-115.48054	Owyhee	ID	July 9	7	13.7	159	20
Almo	42.15626	-113.72014	Cassia	ID	July 10	8	8.3	50	32
Clear	41.95195	-113.32645	Box Elder	UT	July 11	8	5.9	85	30
Phelan	45.16683	-114.15586	Lemhi	ID	July 13	6	7.4	37	38
Dry	44.12386	-113.57582	Butte	ID	July 20 & 21	8	9.9	295	31
Kenney	45.03917	-113.62498	Lemhi	ID	July 22 & 23	8	9.1	56	40
East Fork Daly	45.28670	-114.03542	Lemhi	ID	July 23 & 24	8	7.4	30	27

Table 2.1 (cont.)

Moose	45.31057	-114.04138	Lemhi	ID	July 24 & 25	11	9.2	24	37
East Boulder	45.35387	-114.15075	Lemhi	ID	July 26 & 27	6	9.5	25	31
Fox	43.87089	-113.92258	Custer	ID	July 29 & 30	8	9.5	164	28
Cabin	43.82749	-113.84859	Custer	ID	July 30	8	12.9	236	33
East Fork Big Lost	43.82555	-113.84948	Custer	ID	July 30	8	11.4	201	37
Bear	43.67616	-113.69364	Custer	ID	July 31	8	11.4	153	35
Starhope	43.77151	-113.93269	Custer	ID	July 31	8	8.2	96	35
Horsethief	43.62678	-113.72775	Butte	ID	August 1	8	9.3	253	35
Bearskin	44.40842	-115.49325	Valley	ID	August 3 & 4	7	8.3	38	30
Trapper	44.79614	-115.48531	Valley	ID	August 4 & 5	8	6.9	58	39
Little Goose	44.96196	-116.16825	Adams	ID	August 6	8	12.1	70	38
East Branch Weiser	45.02185	-116.43179	Adams	ID	August 7	8	15.2	121	32
Lost	44.88536	-116.43404	Adams	ID	August 8 & 9	8	18.4	54	27
Boardman	43.60799	-114.94109	Camas	ID	August 17 & 18	6	9.4	79	36
Leggit	43.81568	-115.04781	Elmore	ID	August 19	6	6.3	67	21
Decker	43.76836	-115.14054	Elmore	ID	August 20	6	10.5	70	31
Little Queens	43.90210	-115.18638	Elmore	ID	August 21	6	8.2	44	23
Queens	43.85534	-115.13880	Elmore	ID	August 22	6	10.0	24	31
Black Warrior	43.84070	-115.25607	Elmore	ID	August 23	6	7.7	68	41
Swanholm	43.82609	-115.36035	Elmore	ID	August 24	6	10.2	101	43
Big	41.58457	-111.30839	Rich	UT	September 10	8	10.9	409	24
Otter	41.71523	-111.25610	Rich	UT	September 11	8	8.8	375	30
Rock	41.60867	-111.58630	Cache	UT	September 12	6	10.6	375	24

Table 2.2. Macroinvertebrate densities (per square meter) and frequencies across the 332 sampled patches. The * denotes the twelve taxa included in the analysis to assess interactive effects of habitat attributes on taxon densities.

Taxon	Taxon Type	Total Density (per m^2)	Mean Density (per m^2)	Median Density (per m^2)	Maximum Density (per m^2)	Frequency (number of patches in which taxon was found)
Baetis	Mayfly	160,739	484	174	10,261	279
Cinygmula	Mayfly	52,222	157	49	1,826	199
Oligochaeta	Annelid worm	115,853	349	33	32,116	185
**Simulium*	Black fly	19,979	60	0	1,449	150
Optioservus	Riffle beetle	180,146	543	0	16,957	150
**Optioservus quadrimaculatus*	Riffle beetle	57,332	173	0	3,391	134
Lebertia	Mite	15,476	47	0	1,044	134
Epeorus	Mayfly	33,096	100	0	1,696	131
Heterlimnius corpulentus	Riffle beetle	34,807	105	0	2,348	130
Diphetor hageni	Mayfly	18,194	55	0	1,304	129
Rhyacophila	Caddis fly	9,562	29	0	696	127
Pisidium	Clam	71,770	216	0	15,826	125
**Rhyacophila brunnea vemna group*	Caddis fly	8,748	26	0	1,044	125
Perlodidae	Stonefly	12,690	38	0	609	125
**Ameletus*	Mayfly	10,938	33	0	1,065	109
Drunella coloradensis flavilinea	Mayfly	11,897	36	0	1,935	106
Micrasema	Caddis fly	26,625	80	0	2,870	104
Hexatoma	Crane fly	6,705	20	0	522	99
**Torrenticola*	Mite	11,014	33	0	565	90
Sweltsa	Stonefly	8,402	25	0	565	89
Turbellaria	Flatworm	15,579	47	0	1,413	83
Probezzia	Midge	6,824	21	0	609	81
Paraleptophlebia	Mayfly	13,769	41	0	1,304	80

Table 2.2 (cont.)

Sperchon	Mite	9,572	29	0	957	78
**Drunella doddsii*	Mayfly	27,626	83	0	3,217	75
Serratella	Mayfly	7,121	21	0	609	72
**Suwallia*	Stonefly	8,977	27	0	826	66
Nemata	Nematode	11,581	35	0	2,783	60
Cleptelmis addenda	Riffle beetle	5,017	15	0	435	59
Protzia	Mite	6,772	20	0	696	58
Antocha monticola	Crane fly	7,376	22	0	891	57
Glossosoma	Caddis fly	5,829	18	0	1,435	55
Zaitzevia	Riffle beetle	5,597	17	0	1,174	55
Capniidae	Stonefly	5,737	17	0	880	54
Drunella	Mayfly	7,372	22	0	754	54
Neophylax	Caddis fly	3,820	12	0	348	53
Zapada	Stonefly	6,348	19	0	1,087	52
Ephemerella inermis dorothea	Mayfly	31,562	95	0	5,913	44
Rhithrogena	Mayfly	4,576	14	0	544	42
Agapetus	Caddis fly	15,134	46	0	2,087	42
Hydropsyche	Caddis fly	4,039	12	0	522	39
Dicranota	Crane fly	2,547	8	0	522	36
**Yoraperla*	Stonefly	16,407	49	0	9,652	36
Glutops	Fly	2,271	7	0	402	34
Malenka	Stonefly	12,097	36	0	2,609	34
Hydroptilidae	Caddis fly	7,260	22	0	3,130	33
Arctopsyche grandis	Caddis fly	3,110	9	0	696	33
Heterlimnius	Riffle beetle	5,962	18	0	913	32
Brachycentrus americanus	Caddis fly	6,219	19	0	3,609	32
**Physa*	Snail	6,876	21	0	2,783	31
**Acentrella*	Mayfly	2,034	6	0	304	30

Table 2.2 (cont.)

Lepidostoma	Caddis fly	1,553	5	0	261	29
Hesperoperla pacifica	Stonefly	1,163	4	0	87	27
Rhyacophila sibirica group B	Caddis fly	1,452	4	0	174	26
Narpus concolor	Riffle beetle	1,953	6	0	217	26
Centroptilum	Mayfly	12,004	36	0	2,667	26
Serratella tibialis	Mayfly	2,857	9	0	652	26
**Dicosmoecus*	Caddis fly	631	2	0	65	24
**Limnophila*	Crane fly	637	2	0	87	23
Neoplasta	Dance fly	4,406	13	0	1,391	23
Leuctridae	Stonefly	773	2	0	87	22
Pericoma	Moth fly	1,507	5	0	457	21
Apatania	Caddis fly	3,575	11	0	870	20

Table 2.3. Bins used in the electivity index and observed:expected ratio analyses. Each of the environmental gradients was divided into five bins, where n is the number of samples in each bin. Temp = temperature, Cond = conductivity, Vel = velocity, D50 = median particle size, PCLT2 = percent fine sediment < 2mm.

Bins	Temp (°C)	n	Cond (µS/cm)	n	Vel (m/s)	n	D50 (mm)	n	PCLT2 (%)	n
1	< 7	22	< 50	50	< 0.2	80	< 10	35	< 4	137
2	≥ 7-9	99	≥ 50-150	125	≥ 0.2-0.4	123	≥ 10-30	83	≥ 4-8	72
3	≥ 9-11	118	≥ 150-250	63	≥ 0.4-0.6	76	≥ 30-50	98	≥ 8-16	71
4	≥ 11-13	54	≥ 250-350	29	≥ 0.6-0.8	32	≥ 50-70	83	≥ 16-32	39
5	≥ 13	39	≥ 350	65	≥ 0.8	21	≥ 70	33	≥ 32	13

Table 2.4. Descriptions of each of the five environmental gradients used to estimate macroinvertebrate tolerance values, including units and descriptive statistics.

Abbreviation	Description	Units	Median	Minimum	Maximum
Temp	Temperature of each stream, standardized to July 15 at 11:30 a.m.	°C	9.5	5.9	18.4
Cond	Conductivity of each stream	µS/cm	108	24	542
Vel	Water velocity occurring in each stream patch	m/s	0.29	0.0	1.6
d50	Median particle size of the substrate in each stream patch	mm	40.8	0.6	87.7
PCLT2	Percentage of each patch's substrate made up of fine sediments (< 2 mm)	%	5.2	0.0	97.1

Table 2.5. Spearman's rank correlation coefficients between environmental variables for both reach-scale (n = 45) and patch-scale (n = 332) data. Temp = temperature, Cond = conductivity, Vel = velocity, D50 = median particle size, and PCLT2 = percent fine sediment. The * indicates the correlation is significant at a Bonferroni-adjusted alpha of 0.005.

		Temp	Cond	Vel	D50
Reach-scale data	Cond	0.43*			
	Vel	-0.16	0.13		
	D50	-0.19	-0.32	0.06	
	PCLT2	0.11	0.26	0.07	-0.76*
Patch-scale data	Cond	0.43*			
	Vel	-0.04	0.15		
	D50	-0.16*	-0.22*	0.13	
	PCLT2	0.10	0.17*	-0.08	-0.79*

Table 2.6. Spearman's rank correlation coefficients comparing the six types of estimates for each of the five environmental variables across all taxa (n = 63). WA = weighted averages, DPeak = peak of the density, O:E = observed:expected ratios, SI = selectivity indices, PPeak = peak of the presence curve, and 75^{th} = 75^{th} percentiles. The * indicates the correlation is significant at a Bonferroni-adjusted alpha of 0.003.

		WA	DPeak	O:E	SI	PPeak
Temperature	DPeak	0.70*				
	O:E	0.85*	0.62*			
	SI	0.54*	0.32	0.49*		
	PPeak	0.53*	0.32	0.41*	0.48*	
	75th	0.79*	0.37*	0.64*	0.66*	0.70*
Conductivity	DPeak	0.77*				
	O:E	0.82*	0.64*			
	SI	0.43*	0.33	0.44*		
	PPeak	0.39*	0.39*	0.36	0.29	
	75th	0.82*	0.57*	0.75*	0.59*	0.41*
Velocity	DPeak	0.48*				
	O:E	0.81*	0.45*			
	SI	0.60*	0.25	0.41*		
	PPeak	0.25	0.03	0.23	0.23	
	75th	0.67*	0.04	0.50*	0.50*	0.24
D50	DPeak	0.70*				
	O:E	0.78*	0.61*			
	SI	0.47*	0.23	0.50*		
	PPeak	0.46*	0.24	0.21	0.31	
	75th	0.60*	0.31	0.43*	0.38*	0.42*
Percent Fines	DPeak	0.70*				
	O:E	0.85*	0.70*			
	SI	0.66*	0.35	0.52*		
	PPeak	0.00	0.07	0.04	-0.11	
	75th	0.71*	0.27	0.54*	0.58*	0.13

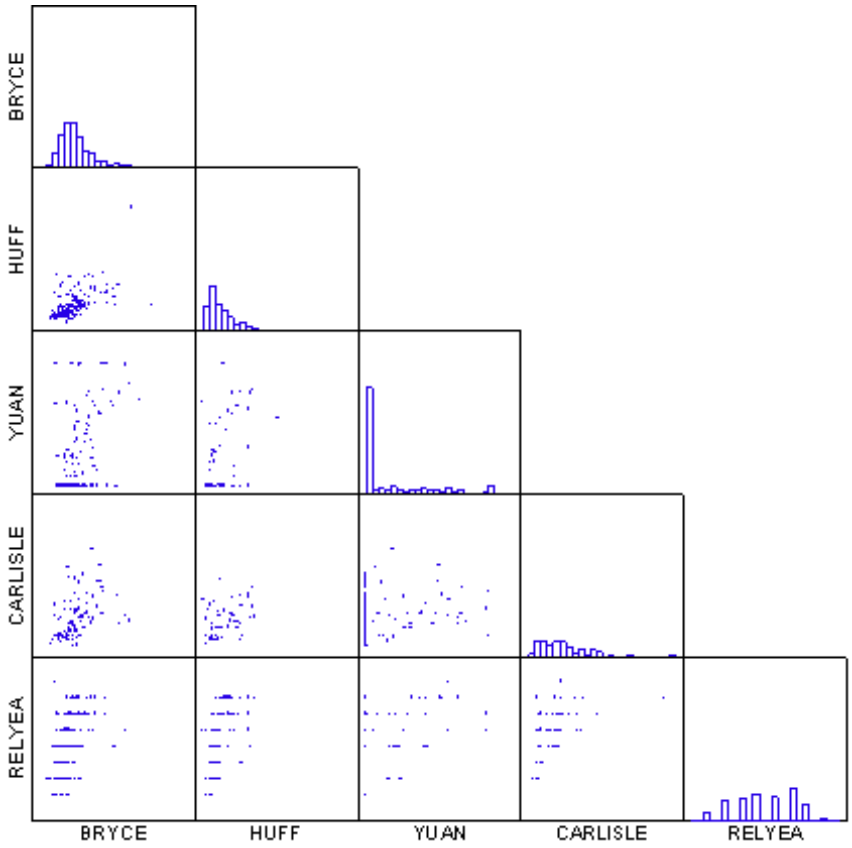

Fig. 2.1. Scatterplot matrix comparing tolerance values for percent fine sediment for several different taxa across five studies (Bryce et al. 2010, Carlisle et al. 2007, Relyea 2007, Huff et al. 2008, Yuan 2006). Each data point represents a taxon for which an optimum was estimated in at least two studies.

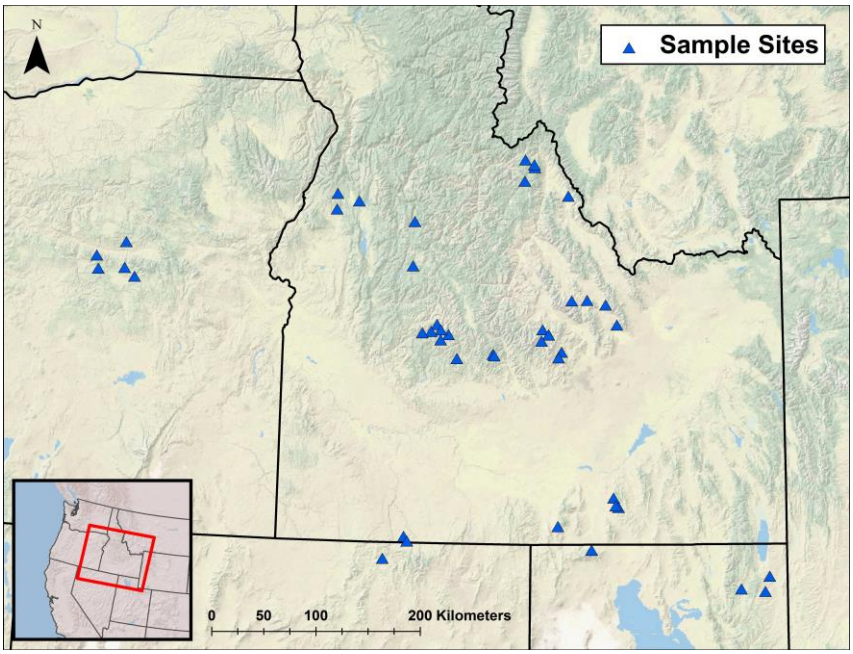

Fig. 2.2. Locations of the 45 sampled streams (indicated by blue triangles) in four states in the western U.S.

	Low Temp	High Temp
Low Cond	<9.5°C <70 µS/cm ---------- 82 (11)	≥ 9.5°C <70 µS/cm ---------- 28 (4)
Med Cond	<9.5°C 70 to 225 µS/cm ---------- 45 (6)	≥ 9.5°C 70 to 225 µS/cm ---------- 67 (9)
High Cond	<9.5°C >225 µS/cm ---------- 35 (5)	≥ 9.5°C >225 µS/cm ---------- 75 (10)

Fig. 2.3. Six temperature and conductivity bins representing the environmental space that existed across the 45 sites. The temperature and conductivity values of each bin are given above the dashed lines. The number of patches and streams in each bin are given below each dashed line [patches (streams)].

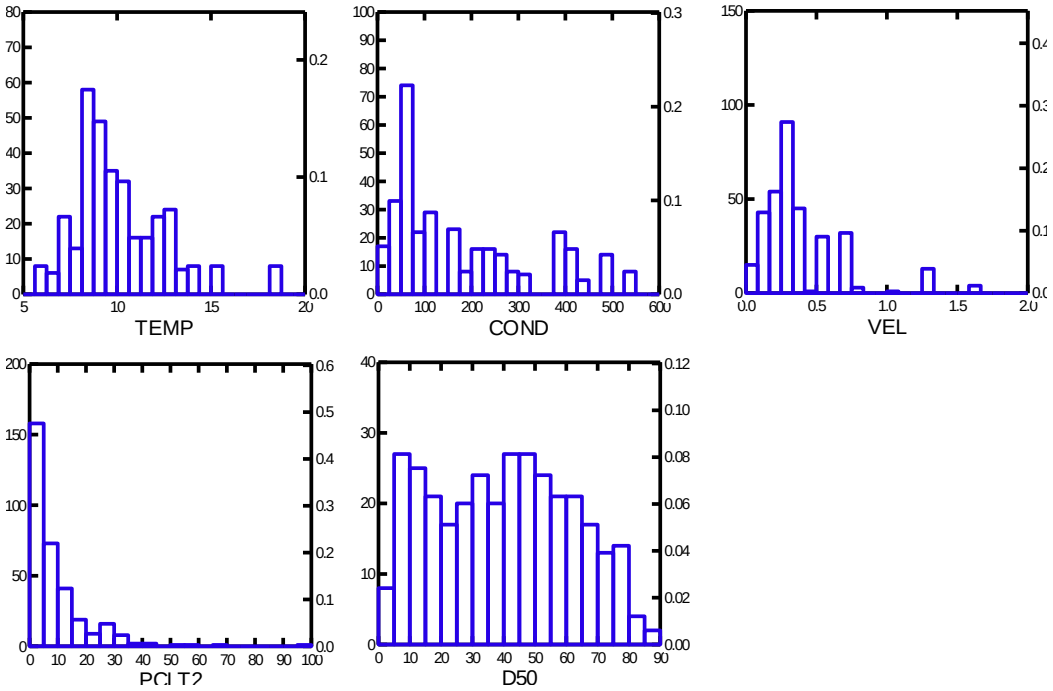

Fig. 2.4. Distribution of values for the five environmental factors across the 332 samples.

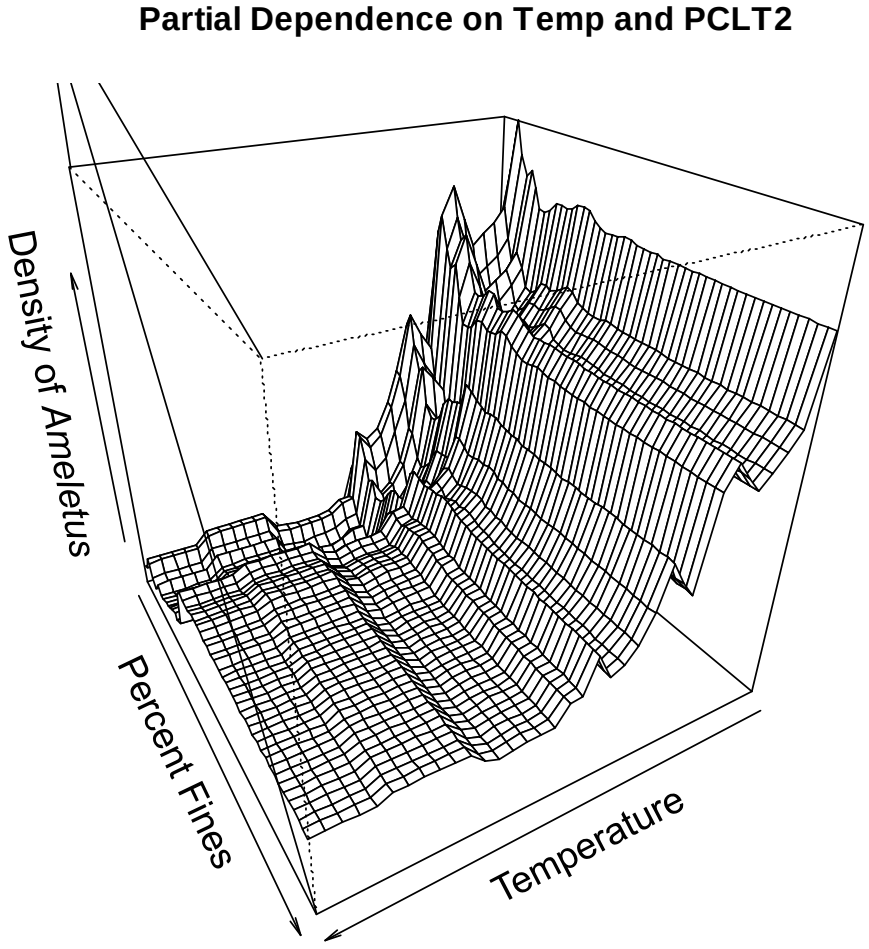

Fig. 2.5. Bivariate partial dependence plot showing the logged density response of *Ameletus* to both temperature and percent fine sediment.

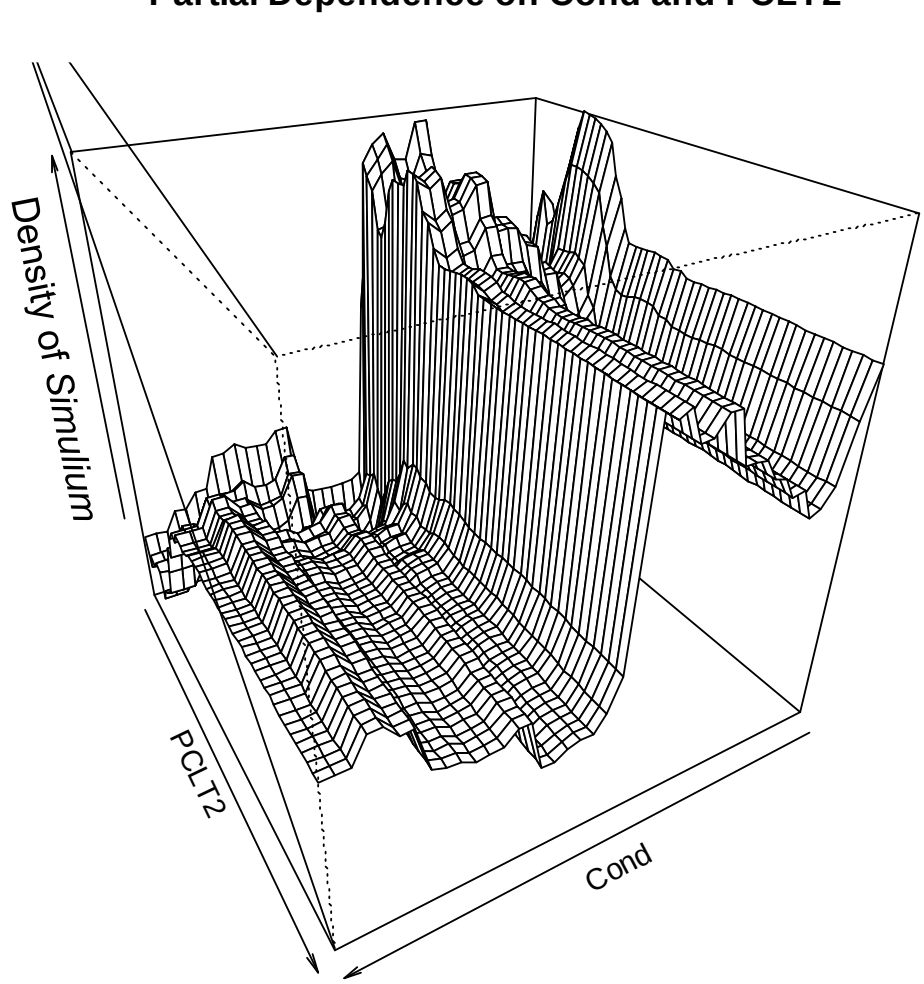

Fig. 2.6. Bivariate partial dependence plot showing the logged density response of *Simulium* to both conductivity and percent fine sediment.

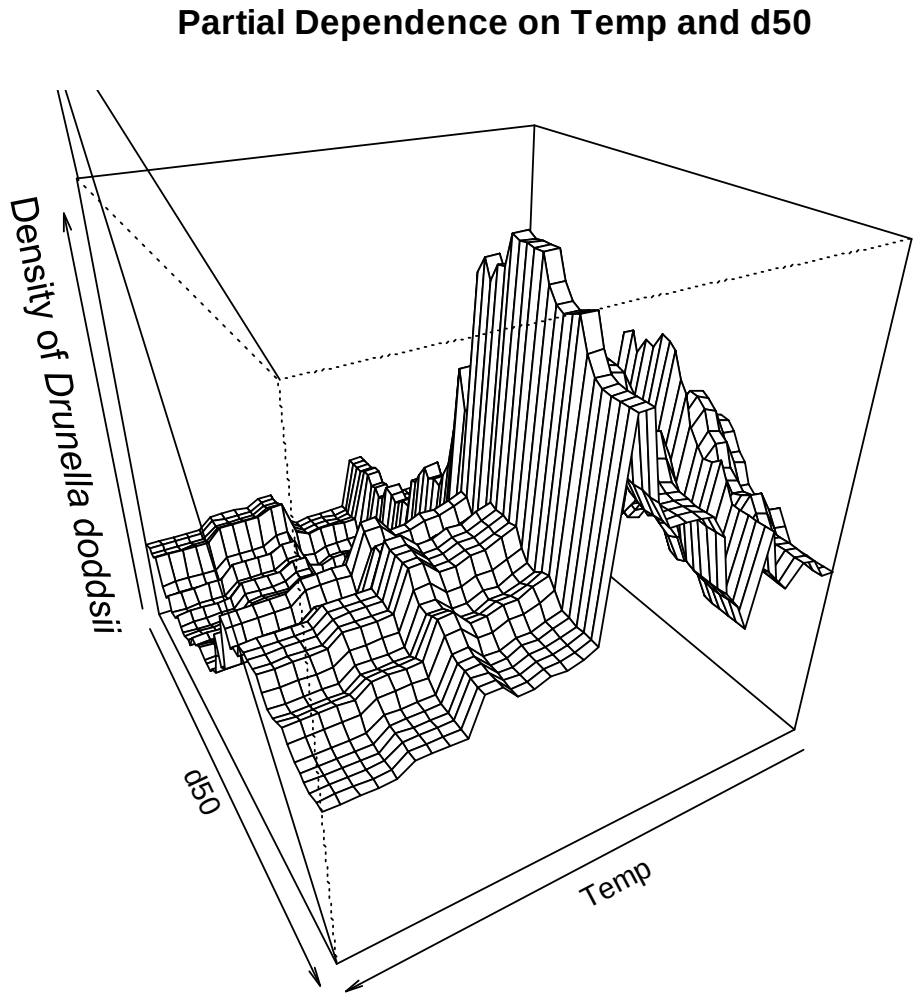

Fig. 2.7. Bivariate partial dependence plot showing the logged density response of *Drunella doddsii* to both temperature and median particle size.

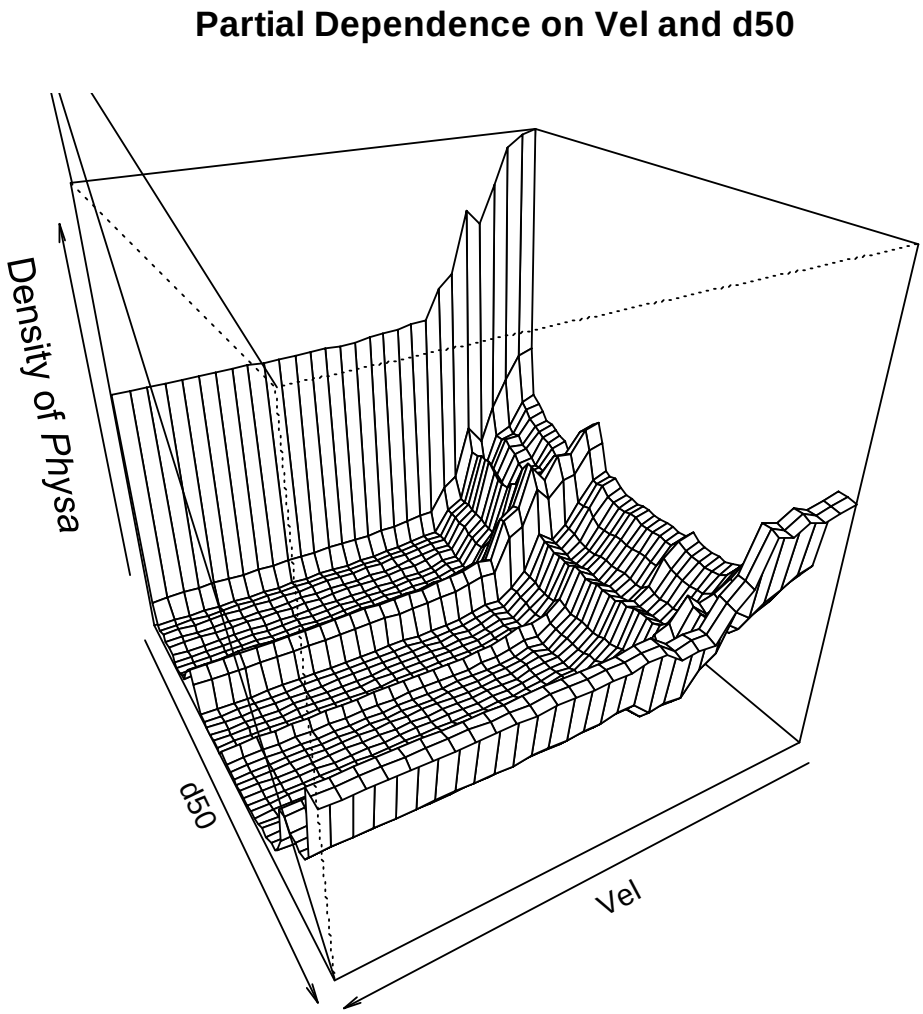

Fig. 2.8. Bivariate partial dependence plot showing the logged density response of *Physa* to both velocity and median particle size.

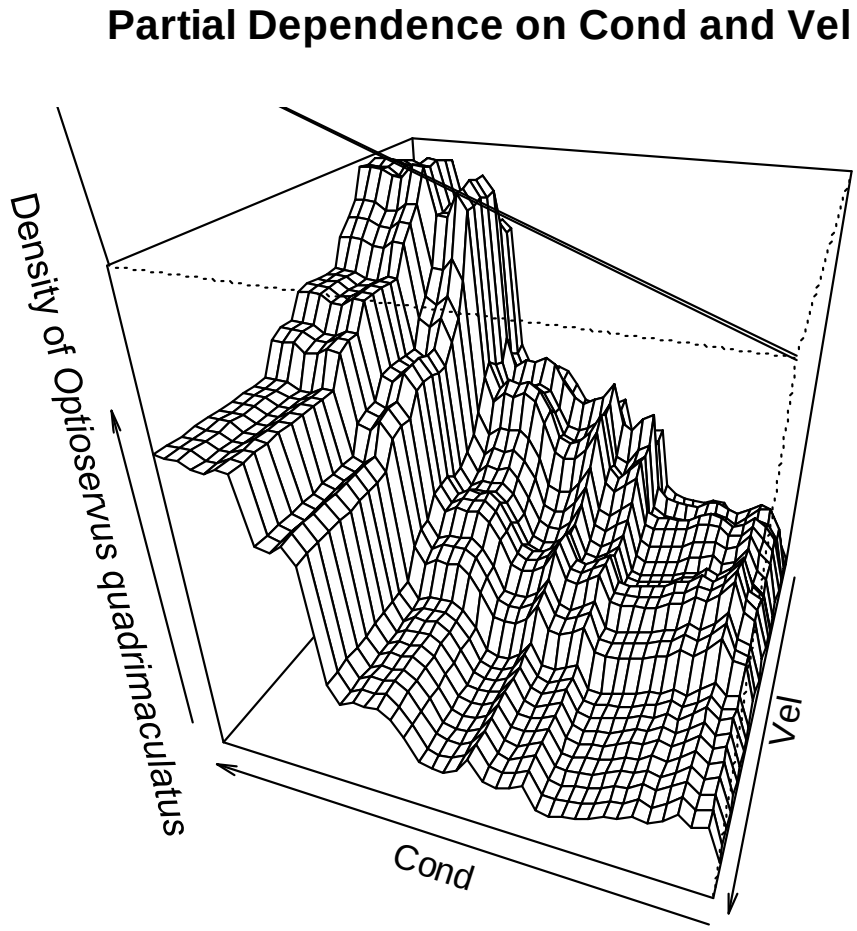

Fig. 2.9. Bivariate partial dependence plot showing the logged density response of *Optioservus quadrimaculatus* to both conductivity and velocity.

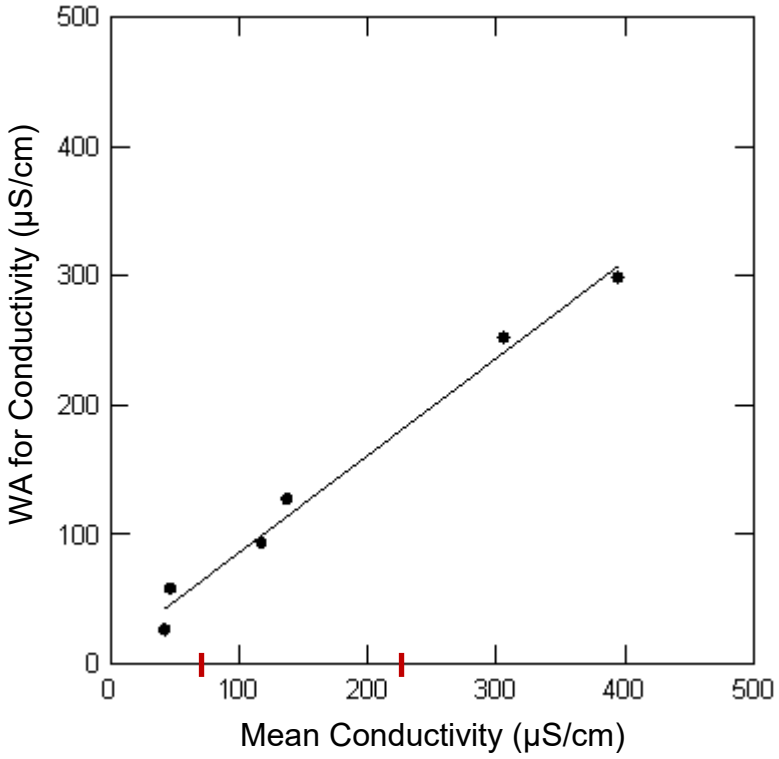

Fig. 2.10. Scatterplot demonstrating the relationship between the weighted average conductivity preference estimates for *Epeorus* and the mean conductivity for each of the six bins. The red lines on the x-axis indicate the cut-off values between the three conductivity bins (70 and 225 µS/cm).

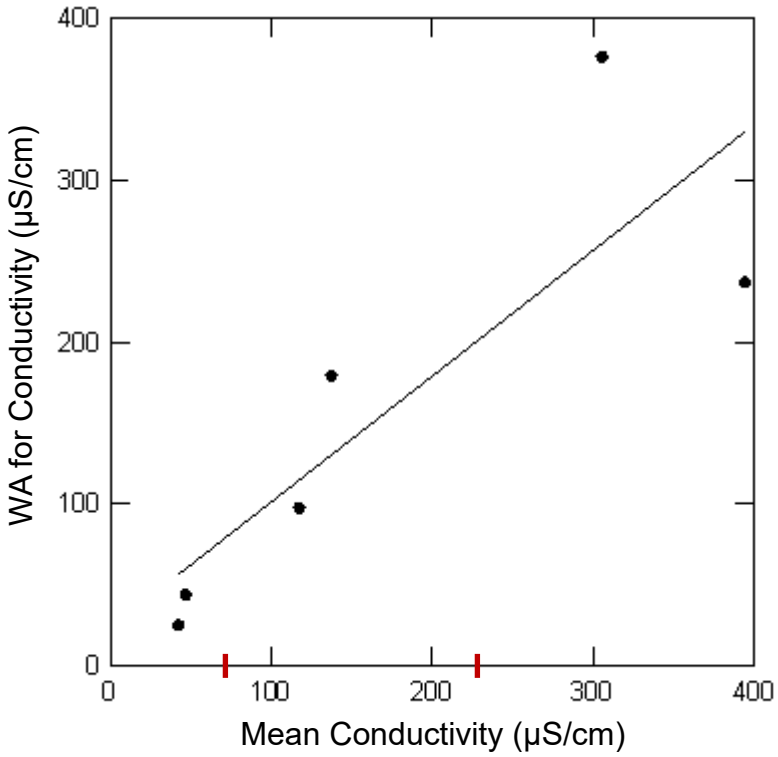

Fig. 2.11. Scatterplot demonstrating the relationship between the weighted average conductivity preference estimates for *Acentrella* and the mean conductivity for each of the six bins. The red lines on the x-axis indicate the cut-off values between the three conductivity bins (70 and 225 µS/cm).

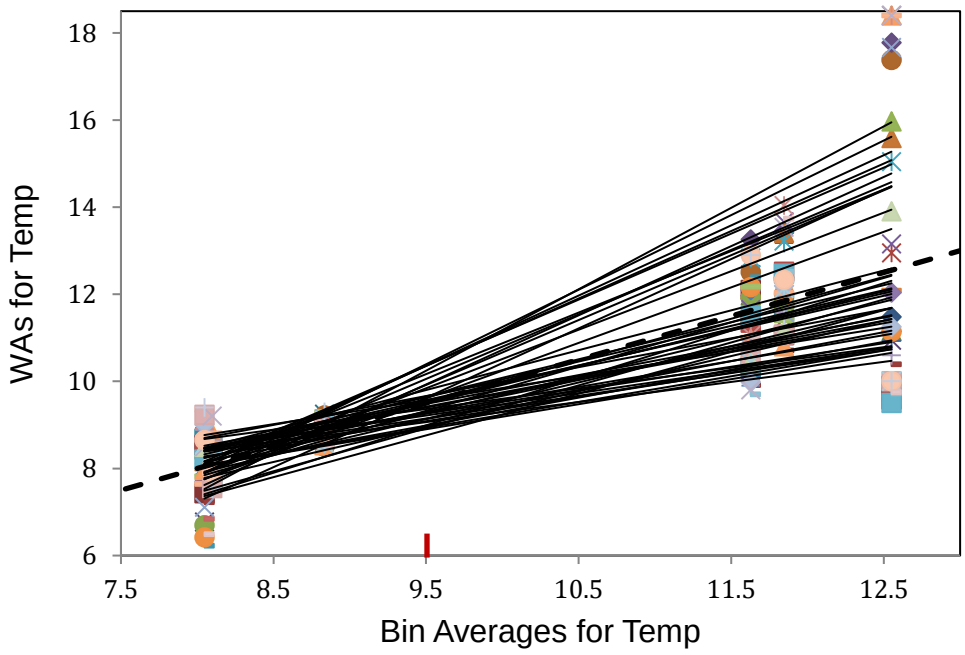

Fig. 2.12a. Scatterplot demonstrating the relationships between weighted average temperature preference estimates and mean temperatures for each of the six bins for 43 of the 63 taxa. These 43 taxa occurred in streams in all six temperature/conductivity bins. (See Table A.1 in the Appendix for a list of taxa included/excluded in this figure. Symbols are not tied to specific taxa.) The dashed line represents the 1:1 relationship between the weighted averages and the bin averages; the solid lines are the trend-lines for each taxon. The red line on the x-axis indicates the cut-off value between the two temperature bins (9.5°C).

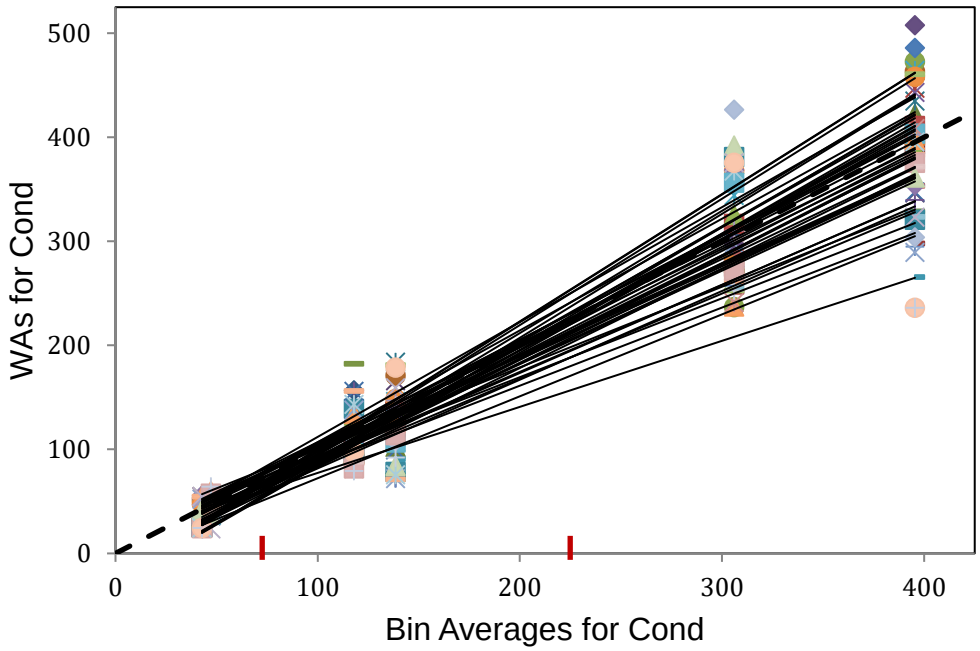

Fig. 2.12b. Scatterplot demonstrating the relationships between weighted average conductivity preference estimates and mean conductivities for each of the six bins for 43 of the 63 taxa. These 43 taxa occurred in streams in all six temperature/conductivity bins. (See Table A.1 in the Appendix for a list of taxa included/excluded in this figure. Symbols are not tied to specific taxa.) The dashed line represents the 1:1 relationship between the weighted averages and the bin averages; the solid lines are the trend-lines for each taxon. The red lines on the x-axis indicate the cut-off values between the three conductivity bins (70 and 225 µS/cm).

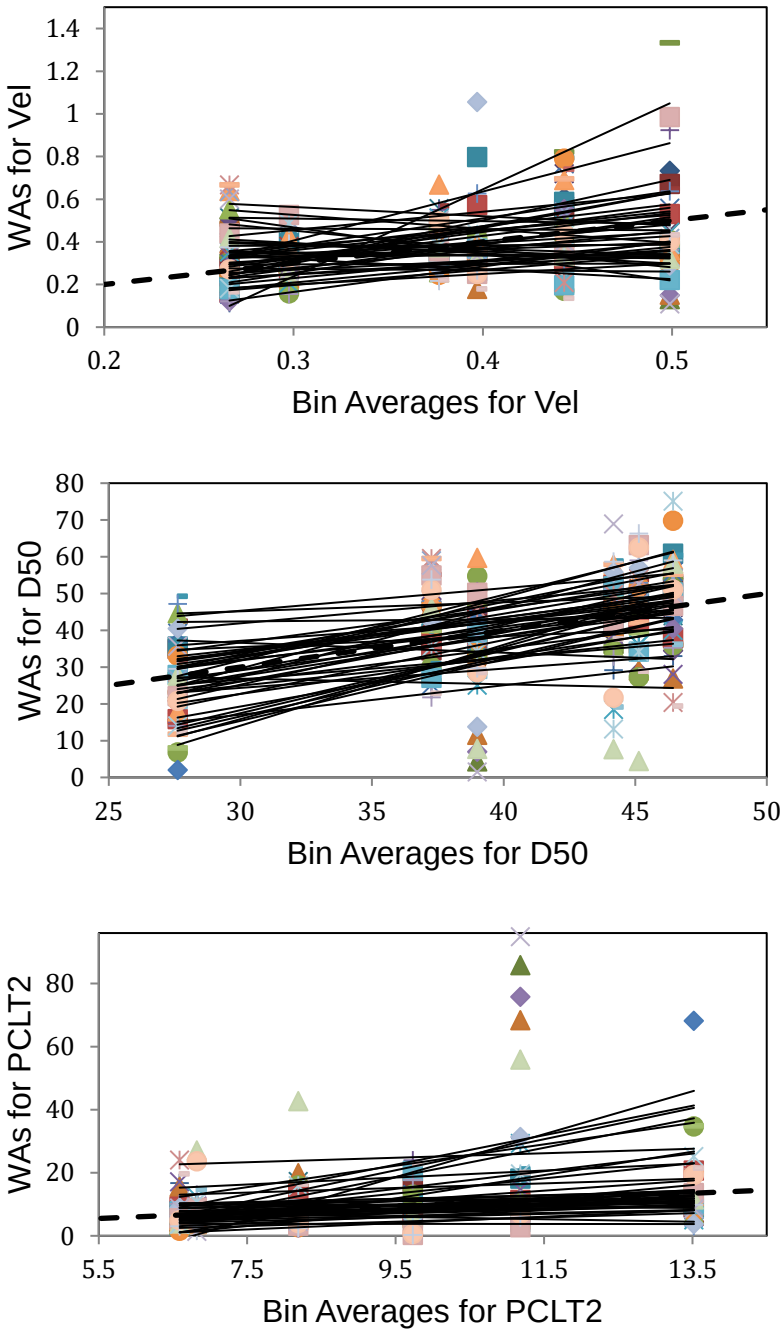

Fig. 2.13. Scatterplots demonstrating the relationships between velocity (m/s), median particle size (mm), and percent fines (%) weighted average preference estimates and means for each of the six bins for 43 of the 63 taxa. These 43 taxa occurred in streams in all six temperature/conductivity bins. (See Table A.1 in the Appendix for a list of taxa included/excluded in this figure. Symbols are not tied to specific taxa.) The dashed line represents the 1:1 relationship between the weighted averages and the bin averages; the solid lines are the trend-lines for each taxon.

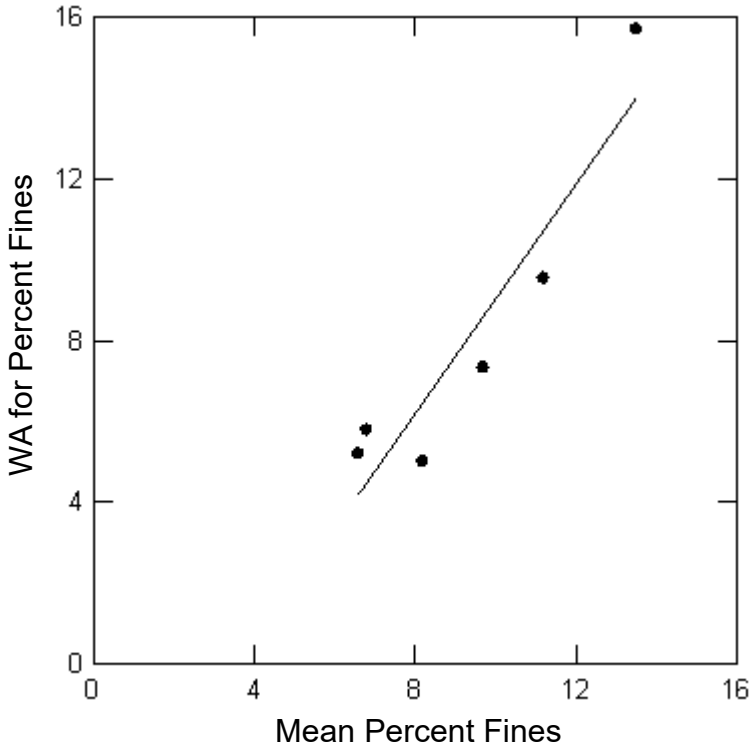

Fig. 2.14. Scatterplot demonstrating the relationship between the weighted average percent fines preference estimates for *Serratella* and the mean percent fines for each of the six temperature/conductivity bins.

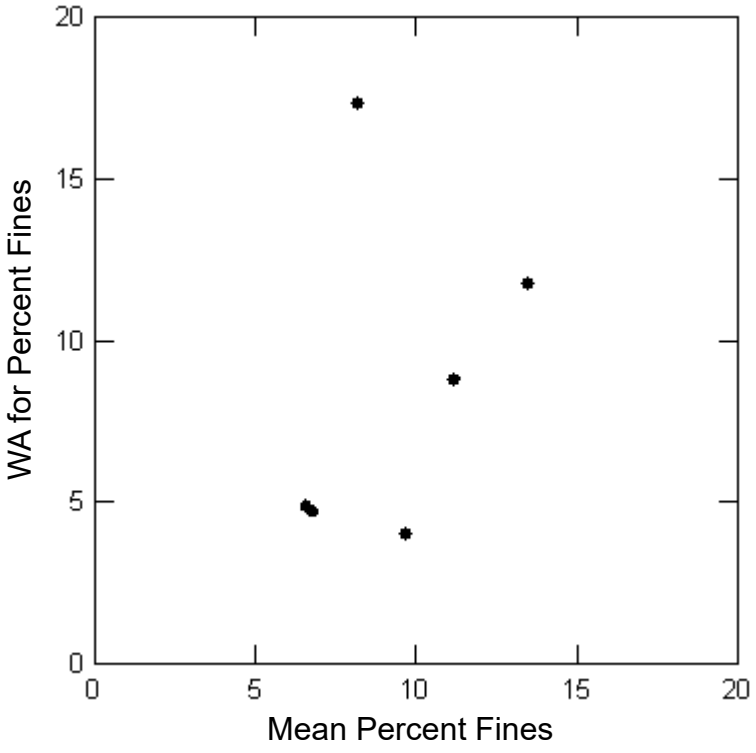

Fig. 2.15. Scatterplot demonstrating the lack of relationship between the weighted average percent fines preference estimates for *Optioservus* and the mean percent fines for each of the six temperature/conductivity bins.

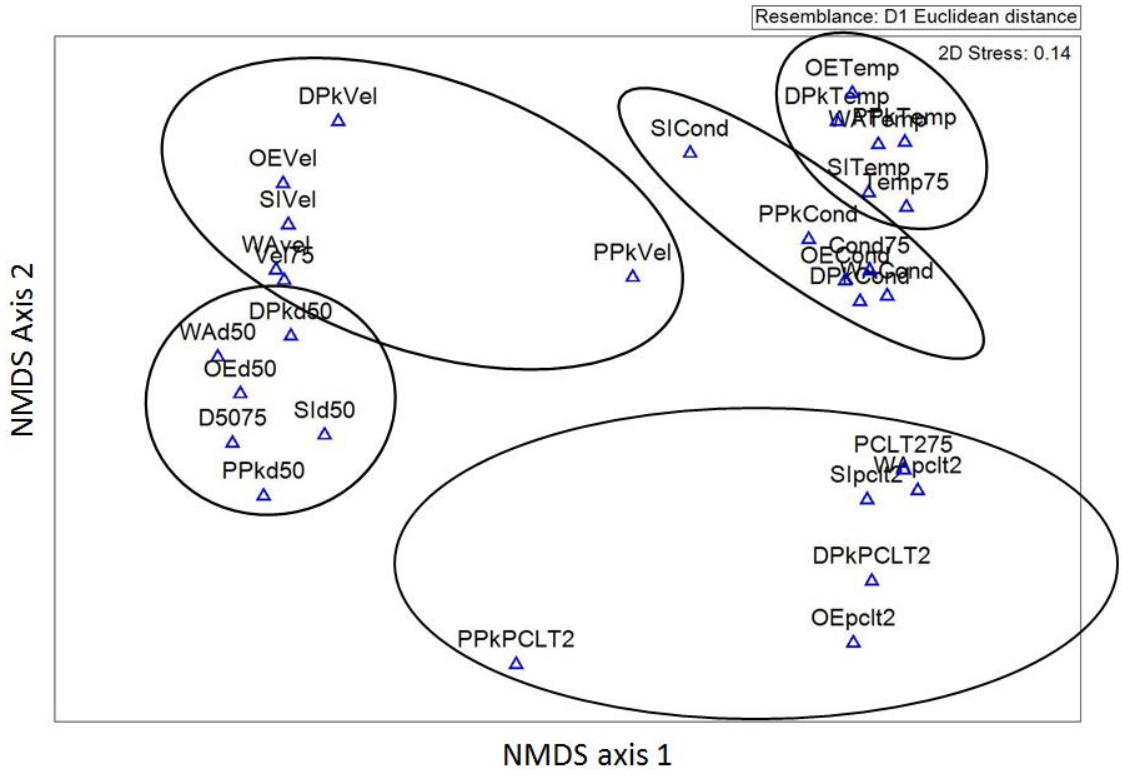

Fig. 2.16. NMDS ordination (Euclidean distance measures, two-dimensional, stress = 0.14) showing the proximity of different tolerance values to one another in environmental tolerance space. The circles were manually drawn to further demonstrate this separation.

CHAPTER 3

THE FUTURE UTILITY OF STRESSOR-SPECIFIC MACROINVERTEBRATE TOLERANCE VALUES

Biotic indices are common tools in bioassessments. These indices, which are "single numbers that summarize complex biological data and reflect water quality" (Stark 1998), weight tolerance metrics by macroinvertebrate abundances to assign a score to a stream. Existing tolerance-based biotic indices include the Hilsenhoff Biotic Index (United States), the Macroinvertebrate Community Index (New Zealand), the SIGNAL index (Australia), and the Biological Monitoring Working Party score (Britain). Multimetric indices incorporate various metrics describing assemblage composition (e.g., Simpson's diversity, Shannon's diversity, taxon richness), ecology (e.g., relative abundance of predators, % filterers, number of functional trait niches), and taxon or assemblage tolerance to stressors (e.g., % intolerant EPT taxa, relative abundance of tolerant taxa, % tolerant Chironomids) (Karr and Chu 1999; Weigel and Dimick 2011). In a multimetric bioassessment approach, each metric is assigned a value for each site, and then those values are aggregated into the overall biotic index score (Chessman and McEvoy 1998; Norris and Hawkins 2000; Weigel and Dimick 2011). Stream impairment is inferred at a site if the index score falls below some predefined threshold. In a multivariate approach, the composition of biota observed at a test site is compared to the biota that a predictive model expects based on reference conditions. Stream impairment is inferred if the observed/expected ratio falls outside prediction error (Norris and Hawkins 2000).

In the last couple of decades, researchers have begun to add biological traits to the suite of metrics used in bioassessments (Culp et al. 2011). These biological traits can include those representing life history (e.g., body size, reproductive strategy, life cycle), resistance and resilience (e.g., locomotion, body form, dissemination potential), or general biology (e.g., feeding habit, respiration strategy) (Statzner and Beche 2010, Monaghan and Soares 2012, Lange et al. 2014). The use of traits to determine causes of stream impairment originates from the understanding that environmental conditions (or "habitat filters") influence the expression of traits in taxa at multiple hierarchical spatial scales (Townsend et al. 1997, Poff 1997, Culp et al. 2011). In the case of an impaired stream, a stressor (e.g., an increase in deposited sediment) is one "filter" through which only those taxa with certain traits (e.g., burrowing habit) can pass.

The use of biological traits to diagnose causes of impairment has several potential advantages over the use of community composition metrics. Biological traits tend to be more stable across large spatial scales, temporal scales, and geographic locations than species composition (Culp et al. 2011, Lange et al. 2014). For example, Pollard and Yuan (2010) found that the relationship between richness of clingers (a trait) and increased stream sediments was much more consistent across spatial scales and disparate locations than the relationship between EPT richness (a community composition metric) and increased sediments. Also, rather than operating under the assumption that the abundance or presence of taxa directly reflects some, typically unknown, aspect of the quality of the habitat (as community composition metrics do; Van Horne 1983, Karr and

Chu 1999), biological traits should imply a specific mechanistic link between biological responses and environmental conditions (Statzner and Beche 2010, Culp et al. 2011).

Another advantage of using biological traits in bioassessments is that they have the potential to disentangle the effects of multiple stressors (Beketov et al. 2009, Statzner and Beche 2010, Wooster et al. 2012). For example, Wooster et al. (2012) found that resilience traits (e.g., rates of dispersal and drift) were able to determine that scouring flows had a greater impact on stream macroinvertebrate assemblages than water withdrawals, whereas community composition metrics were unable to make this distinction. However, using traits to assess the cause of impairment requires *a priori* predictions of that cause and a careful selection of traits that could provide a mechanistic explanation for the actions of the stressor (Statzner and Beche 2010). An example described by Statzner and Beche (2010) illustrates the process of trait selection. In the example, the authors try to distinguish between two causes of oxygen deficiency: organic pollution and cold-water releases from the oxygen-poor hypolimnion behind a dam. If organic pollution is the cause of oxygen deficiency, then respiratory traits such as air-breathing alone should increase. If air-breathing increases in combination with a temperature-dependent trait (e.g., decrease in mussel shell growth), then cold, oxygen-poor reservoir water is more likely to be the cause of oxygen deficiency.

A recent development in bioassessment research is the building of biotic indices from multiple biological and physiological trait metrics. Monaghan and Soares (2012) revised the BMWP (Biological Monitoring Working Party) index and included seven different types of traits (respiration, habit, trophic class, flow preference, thermal

preference, size, and activity level) to describe sensitivity to organic pollution. These types of traits were included because organic pollution tends to cause a decrease in the level of dissolved oxygen and an increase in turbidity, sedimentation, and nutrient concentration. The deposition of sediment changes benthic habitats and can cause an increase in heterotrophic production (Monaghan and Soares 2012). The authors compared BMWP index scores to the trait index scores, and trait index scores did not prove to be more sensitive to organic pollution. However, the authors noted that including other traits (e.g., presence of hemoglobin, facultative surface breathing, and food type) could improve the sensitivity of their traits-based index. Another trait-based biotic index is the SPEAR (SPEcies At Risk) index (Beketov et al., 2009, Schafer et al. 2011), which was developed to detect the effects of pesticides (Beketov et al. 2009); however its use has been adapted to detect the effects of salinity (Schafer et al. 2011). The $SPEAR_{pesticide}$ index incorporates four traits: generation time, migration abilities, presence of aquatic stages during maximum pesticide usage, and sensitivity to organic toxicants (Beketov et al. 2009). Four different traits were included in the $SPEAR_{salinity}$ index: reproduction type, food source, respiration mode, and physiological sensitivity to salinity (Schafer et al. 2011). The combination of physiological and biological traits was successful in identifying taxa sensitive to pesticides, whereas the physiological trait alone best identified taxa at risk from salinity. This difference is likely linked to the type of disturbance being considered. Pesticides typically occur as a pulse disturbance, whereas salinity is more likely to be a press disturbance. Therefore, both resistance/resilience traits and physiological traits would determine species at risk from pesticides.

Physiological traits alone would determine species at risk from a continuous exposure to high salinity. Biotic indicators based on *a priori* selected traits, such as the SPEAR index, have so far proven to be a promising tool in bioassessment (Schafer et al. 2011, Monaghan and Soares 2012).

Stressor- and taxon-specific tolerance values, like those estimated for temperature, conductivity, velocity, and substrate in this thesis, are biological trait data and can potentially be used within traits-based indices to determine causes of biological impairment in streams. Tolerance values were originally (and continue to be) estimated for organic pollution and nutrients (Hilsenhoff 1982, Carlisle et al. 2007, Qin et al. 2014), but TVs for many other stressors have been developed as well, including water temperature (Brandt 2001, Carlisle et al. 2007, Cristobal et al. 2014), sediment (Zweig and Rabeni 2001, Huff et al. 2008, Bryce et al. 2010), and conductivity/salinity (Kefford et al. 2003, Carlisle et al. 2007). Chessman and McEvoy (2012) suggest that the absolute tolerance of a taxon to a specific stressor could help to verify or eliminate candidate causes of stream impairment. However, Statzner and Beche (2010) argue that expanding the use of tolerance values in traits-based bioassessments may not be worth the effort (likely years of experimentation) that would be required to estimate TVs for many more stressors and taxa than are currently available. Though it is true that estimating tolerance values for either new taxa or different stressors will require time and resources, estimating them is unlikely to be a wasted effort. Measures of tolerance or preference for different stressors have been successfully incorporated into traits-based bioassessments – e.g. thermal preference and flow preference (Monaghan and Soares 2012), sensitivity to

organic toxicants and physiological sensitivity to salinity (Schafer et al. 2011) – and may better describe characteristics of biota than other available trait types.

It is clear that a traits-based approach to bioassessment may enhance our current understanding of how environmental conditions can influence the structure of biotic communities in streams. However, the universal adoption of a traits-based approach is hampered by a lack of data and a lack of consistency across current databases. Extensive trait databases are available for macroinvertebrates in North America, Europe, New Zealand, and Australia, but these databases are uncommon in other countries/regions (Statzner and Beche 2010, Schafer et al. 2011). Also, trait names, trait descriptions, and taxonomic resolution are not yet consistent across databases (Culp et al. 2011). Traits also tend to fall into a few coarse categories (e.g. "cold" and "warm" categories of thermal preference), rather than along a continuous scale. A robust, universal method for inferring the causes of biological impairment cannot be achieved until species-specific traits that characterize response to different types of stressors are defined for most stream macroinvertebrate taxa. Once this is achieved, however, water quality managers should be able to reliably diagnose the causes of biological impairment in streams and introduce targeted and effective management strategies to combat the causes of impairment and prevent future damage to valuable aquatic resources.

REFERENCES

Beketov, M. A., K. Foit, R. B. Schafer, C. A. Schriever, A. Sacchi, E. Capri, J. Biggs, C. Wells, and M. Liess. 2009. SPEAR indicates pesticide effects in streams – comparative use of species- and family-level biomonitoring data. Environmental Pollution 157:1841-1848.

Brandt, D. 2001. Temperature preferences and tolerances for 137 common Idaho macroinvertebrate taxa. Idaho Department of Environmental Quality. Coeur d'Alene, Idaho.

Bryce, S. A., G. A. Lomnicky, and P. R. Kaufmann. 2010. Protecting sediment-sensitive aquatic species in mountain streams through the application of biologically based streambed sediment criteria. Journal of the North American Benthological Society 29:657-672.

Carlisle, D. M., M. R. Meador, S. R. Moulton II, and P. M. Ruhl. 2007. Estimation and application of indicator values for common macroinvertebrate genera and families of the United States. Ecological Indicators 7:22-33.

Cristobal, E., S. V. Ayuso, A. Justel, and M. Toro. 2014. Robust optima and tolerance ranges of biological indicators: a new method to identify sentinels of global warming. Ecological Research 29:55-68.

Chessman, B. C. and P. K. McEvoy. 1998. Towards diagnostic biotic indices for river macroinvertebrates. Hydrobiologia 364:169-182.

Chessman, B. C. and P. K. McEvoy. 2012. Insights into human impacts on streams from tolerance profiles of macroinvertebrate assemblages. Water, Air, & Soil Pollution 223:1343-1352.

Culp, J. M., D. G. Armanini, M. J. Dunbar, J. M. Orlofske, N. L. Poff, A. I. Pollard, A. G. Yates, and G. C. Hose. 2011. Incorporating traits in aquatic biomonitoring to enhance causal diagnosis and prediction. Integrated Environmental Assessment and Management 7:187-197.

Hilsenhoff, W. L. 1982. Using a biotic index to evaluate water quality in streams. Technical Bulletin No. 132. Department of Natural Resources. Madison, Wisconsin.

Huff, D. D., S. L. Hubler, Y. Pan, and D. L. Drake. 2008. Detecting shifts in macroinvertebrate assemblage requirements: implicating causes of impairment in streams. Oregon Department of Environmental Quality Watershed Assessment. Technical Report: DEQ06-LAB-0068-TR.

Karr, J. R., and E. W. Chu. 1999. Restoring Life in Running Waters: Better Biological Monitoring. Washington, DC: Island. 200 pp.

Kefford, B. J., P. J. Papas, and D. Nugegoda. 2003. Relative salinity tolerance of macroinvertebrates from the Barwon River, Victoria, Australia. Marine and Freshwater Research 54:755-765.

Lange, K., C. R. Townsend, and C. D. Matthaei. 2014. Can biological traits of stream invertebrates help disentangle the effects of multiple stressors in an agricultural catchment? Freshwater Biology 59:2431-2446.

Monaghan, K. A. and A. M. V. M. Soares. 2012. Bringing new knowledge to an old problem: building a biotic index from lotic macroinvertebrate traits. Ecological Indicators 20:213-220.

Norris, R. H. and C. P. Hawkins. 2000. Monitoring river health. Hydrobiologia 435:5-17.

Poff, N. L. 1997. Landscape filters and species traits: towards a mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society 16:391-409.

Pollard, A. I. and L. L. Yuan. 2010. Assessing the consistency of response metrics of the invertebrate benthos: a comparison of trait- and identity-based measures. Freshwater Biology 55:1420-1429.

Qin, C., J. Zhou, Y. Cao, Y. Zhang, R. M. Hughes, and B. Wang. 2014. Quantitative tolerance values for common stream benthic macroinvertebrates in the Yangtze River Delta, Eastern China. Environmental Monitoring and Assessment 186:5883-5895.

Schafer, R. B., B. J. Kefford, L. Metzeling, M. Liess, S. Burgert, R. Marchant, V. Pettigrove, P. Goonan, and D. Nugegoda. 2011. A trait database of stream invertebrates for the ecological risk assessment of single and combined effects of salinity and pesticides in south-east Australia. Science of the Total Environment 409:2055-2063.

Stark, J. D. 1998. SQMCI: a biotic index for freshwater macroinvertebrate coded-abundance data. New Zealand Journal of Marine and Freshwater Research 32:55-66.

Statzner, B. and L. A. Beche. 2010. Can biological invertebrate traits resolve effects of multiple stressors on running water ecosystems? Freshwater Biology 55:80-119.

Townsend, C. R., S. Doledec, and M. R. Scarsbrook. 1997. Species traits in relation to temporal and spatial heterogeneity in streams: a test of habitat templet theory. Freshwater Biology 37:367-387.

Van Horne, B. 1983. Density as a misleading indicator of habitat quality. Journal of Wildlife Management 47:893-901.

Weigel, B. M. and J. J. Dimick. 2011. Development, validation, and application of a macroinvertebrate-based Index of Biotic Integrity for nonwadeable rivers of Wisconsin. Journal of the North American Benthological Society 30:665-679.

Wooster, D. E., S. W. Miller, and S. J. Debano. 2012. An examination of the impact of multiple disturbances on a river system: taxonomic metrics versus biological traits. River Research and Applications 28:1630-1643.

Zweig, L. D. and C. F. Rabeni. 2001. Biomonitoring for deposited sediment using benthic invertebrates: a test on 4 Missouri streams. Journal of the North American Benthological Society 20:643-657.

APPENDICES

Table A.1. The 63 taxa included in this study, sorted both taxonomically and alphabetically. The 20 taxa excluded from the all-taxa temperature/conductivity bin analyses (see Figures 2.12a and 2.12b, 2.13) are marked with *. The three taxa marked with ** are classified taxonomically higher than Order.

Order	Family	Genus	Species
Basommatophora	Physidae	*Physa**	
Coleoptera	Elmidae	*Cleptelmis*	*Cleptelmis addenda*
Coleoptera	Elmidae	*Heterlimnius*	*Heterlimnius corpulentus*
Coleoptera	Elmidae	*Heterlimnius*	
Coleoptera	Elmidae	*Narpus*	*Narpus concolor**
Coleoptera	Elmidae	*Optioservus*	*Optioservus quadrimaculatus*
Coleoptera	Elmidae	*Optioservus*	
Coleoptera	Elmidae	*Zaitzevia*	
Diptera	Ceratopogonidae	*Probezzia*	
Diptera	Empididae	*Neoplasta**	
Diptera	Limoniidae	*Hexatoma*	
Diptera	Limoniidae	*Limnophila*	
Diptera	Pediciidae	*Dicranota*	
Diptera	Pelecorhynchidae	*Glutops**	
Diptera	Psychodidae	*Pericoma**	
Diptera	Simuliidae	*Simulium*	
Diptera	Tipulidae	*Antocha*	*Antocha monticola*
Ephemeroptera	Ameletidae	*Ameletus*	
Ephemeroptera	Baetidae	*Acentrella*	
Ephemeroptera	Baetidae	*Baetis*	
Ephemeroptera	Baetidae	*Centroptilum**	
Ephemeroptera	Baetidae	*Diphetor*	*Diphetor hageni*
Ephemeroptera	Ephemerellidae	*Drunella*	*Drunella coloradensis flavilinea*
Ephemeroptera	Ephemerellidae	*Drunella*	*Drunella doddsii*
Ephemeroptera	Ephemerellidae	*Drunella*	
Ephemeroptera	Ephemerellidae	*Ephemerella*	*Ephemerella inermis dorothea**
Ephemeroptera	Ephemerellidae	*Serratella*	*Serratella tibialis*
Ephemeroptera	Ephemerellidae	*Serratella*	
Ephemeroptera	Heptageniidae	*Cinygmula*	
Ephemeroptera	Heptageniidae	*Epeorus*	
Ephemeroptera	Heptageniidae	*Rhithrogena*	
Ephemeroptera	Leptophlebiidae	*Paraleptophlebia*	
Nemata**			
Oligochaeta**			
Plecoptera	Capniidae		

Table A.1 (cont.)

Plecoptera	Chloroperlidae	*Suwallia*	
Plecoptera	Chloroperlidae	*Sweltsa*	
Plecoptera	Leuctridae*		
Plecoptera	Nemouridae	*Malenka**	
Plecoptera	Nemouridae	*Zapada*	
Plecoptera	Peltoperlidae	*Yoraperla**	
Plecoptera	Perlidae	*Hesperoperla*	*Hesperoperla pacifica**
Plecoptera	Perlodidae		
Trichoptera	Apataniidae	*Apatania**	
Trichoptera	Brachycentridae	*Brachycentrus*	*Brachycentrus americanus**
Trichoptera	Brachycentridae	*Micrasema*	
Trichoptera	Glossosomatidae	*Agapetus**	
Trichoptera	Glossosomatidae	*Glossosoma*	
Trichoptera	Hydropsychidae	*Arctopsyche*	*Arctopsyche grandis**
Trichoptera	Hydropsychidae	*Hydropsyche**	
Trichoptera	Hydroptilidae		
Trichoptera	Lepidostomatidae	*Lepidostoma**	
Trichoptera	Limnephilidae	*Dicosmoecus**	
Trichoptera	Rhyacophilidae	*Rhyacophila*	*Rhyacophila brunnea vemna group*
Trichoptera	Rhyacophilidae	*Rhyacophila*	*Rhyacophila sibirica group B**
Trichoptera	Rhyacophilidae	*Rhyacophila*	
Trichoptera	Uenoidae	*Neophylax*	
Trombidiformes	Hydryphantidae	*Protzia**	
Trombidiformes	Lebertiidae	*Lebertia*	
Trombidiformes	Sperchonidae	*Sperchon*	
Trombidiformes	Torrenticolidae	*Torrenticola*	
Turbellaria**			
Veneroida	Sphaeriidae	*Pisidium*	

Table A.2. Estimates of habitat preference of all 63 taxa for temperature (°C) by six different methods. WA = weighted averaging, DPeak = peak of density curve, O:E = observed:expected ratios, SI = selectivity indices, PPeak = peak of the presence curve, and 75^{th} = 75^{th} percentiles.

Taxon	WA	DPeak	O:E	SI	PPeak	75th
Acentrella	10.9	12	12.1	6.3	8	12.1
Agapetus	10.6	9	12.1	12.1	10.5	11.4
Ameletus	8.6	6	6.3	6.3	8	9.5
Antocha monticola	9.7	9	9.7	6.3	8.5	10.9
Apatania	9.2	8	8.3	6.3	10	10.8
Arctopsyche grandis	9.9	10	9.7	6.3	10.5	11.4
Baetis	10.9	10	13.8	6.3	9	11.4
Brachycentrus americanus	10.1	12	12.1	6.3	8	9.3
Capniidae	9.8	11	9.7	6.3	9.5	10.9
Centroptilum	12.2	12	12.1	13.8	12	13.7
Cinygmula	9.5	13	6.3	6.3	10.5	10
Cleptelmis addenda	10.9	10	13.8	12.1	8	12.1
Dicosmoecus	8.8	8	8.3	6.3	7.5	9.3
Dicranota	9.7	8	8.3	6.3	9	10.5
Diphetor hageni	11.4	13	13.8	13.8	12	12.9
Drunella	9.3	8	8.3	6.3	9.5	9.7
Drunella coloradensis flavilinea	9.8	10	9.7	6.3	8.5	9.7
Drunella doddsii	10.9	12	12.1	6.3	9.5	10.5
Epeorus	9.0	6	6.3	6.3	8.5	10
Ephemerella inermis dorothea	13.3	14	13.8	13.8	11	13
Glossosoma	9.2	7	6.3	6.3	9	10.4
Glutops	8.9	8	8.3	6.3	7.5	9.1
Hesperoperla pacifica	9.9	10	12.1	12.1	10	11.2
Heterlimnius	9.4	13	8.3	8.3	7.5	10.2
Heterlimnius corpulentus	9.6	12	9.7	6.3	8	9.9
Hexatoma	10.2	11	9.7	6.3	9	10.9
Hydropsyche	10.5	13	13.8	13.8	9	11.4
Hydroptilidae	10.8	9	13.8	13.8	14	14.1
Lebertia	9.8	9	9.7	6.3	9	10.6
Lepidostoma	10.0	13	9.7	6.3	10	10.9
Leuctridae	8.7	9	8.3	6.3	7.5	8.9
Limnophila	10.6	14	13.8	6.3	10	11.1
Malenka	12.6	13	12.1	13.8	11	12.9
Micrasema	8.8	7	8.3	6.3	9	10.1
Narpus concolor	9.6	10	8.3	6.3	8	10.4

Table A.2 (cont.)

Nemata	9.7	8	8.3	6.3	9.5	10.9
Neophylax	9.2	8	8.3	6.3	8.5	9.7
Neoplasta	10.7	11	9.7	12.1	8.5	10.9
Oligochaeta	10.1	9	8.3	6.3	10.5	12.1
Optioservus	11.8	11	13.8	13.8	10.5	12.9
Optioservus quadrimaculatus	12.2	14	13.8	13.8	9.5	12.6
Paraleptophlebia	11.9	8	13.8	13.8	12.5	13.4
Pericoma	8.4	8	8.3	9.7	10	10.5
Perlodidae	10.2	8	12.1	6.3	9.5	11
Physa	10.2	9	8.3	13.8	10.5	12.9
Pisidium	11.6	13	13.8	13.8	11.5	12.6
Probezzia	9.5	7	8.3	6.3	7.5	10.6
Protzia	11.0	14	13.8	6.3	10	11.4
Rhithrogena	8.4	7	6.3	6.3	9	9.2
Rhyacophila	9.1	8	8.3	6.3	8	9.9
Rhyacophila brunnea vemna	9.2	9	8.3	6.3	8.5	9.7
Rhyacophila sibirica group B	9.5	12	8.3	6.3	9.5	10.4
Serratella	9.3	6	6.3	6.3	9	11.4
Serratella tibialis	10.9	12	12.1	12.1	10	12.1
Simulium	9.5	6	12.1	6.3	10	12.1
Sperchon	10.8	14	13.8	6.3	12.5	10.9
Suwallia	8.6	9	8.3	6.3	9	9.5
Sweltsa	10.8	12	13.8	6.3	9	10.3
Torrenticola	9.9	10	9.7	6.3	8	10
Turbellaria	10.2	9	13.8	6.3	8.5	10.4
Yoraperla	8.6	8	8.3	6.3	7.5	9.4
Zaitzevia	9.1	7	9.7	13.8	9	10.6
Zapada	8.8	8	8.3	6.3	9.5	10

Table A.3. Estimates of habitat preference of all 63 taxa for conductivity (µS/cm) by six different methods. WA = weighted averaging, DPeak = peak of density curve, O:E = observed:expected ratios, SI = selectivity indices, PPeak = peak of the presence curve, and 75^{th} = 75^{th} percentiles.

Taxon	WA	DPeak	O:E	SI	PPeak	75th
Acentrella	141	175	201	201	100	184
Agapetus	256	325	295	295	250	253
Ameletus	86	100	30	30	25	153
Antocha monticola	202	225	201	201	50	237
Apatania	72	25	30	30	75	151
Arctopsyche grandis	103	100	70	201	75	153
Baetis	252	325	295	295	50	272
Brachycentrus americanus	178	200	201	30	75	218.5
Capniidae	101	50	30	30	25	164
Centroptilum	171	175	201	201	200	184
Cinygmula	177	425	295	295	75	237
Cleptelmis addenda	305	400	423	423	375	375
Dicosmoecus	101	100	70	30	100	108
Dicranota	234	25	295	295	50	324
Diphetor hageni	251	500	423	201	100	375
Drunella	104	25	30	295	25	153
Drunella coloradensis flavilinea	170	100	295	295	50	253
Drunella doddsii	74	75	70	295	75	116
Epeorus	103	75	70	295	50	164
Ephemerella inermis dorothea	507	500	423	295	400	481
Glossosoma	117	75	70	295	75	221
Glutops	256	425	201	201	75	237
Hesperoperla pacifica	213	250	201	295	175	253
Heterlimnius	248	200	423	423	250	394
Heterlimnius corpulentus	149	100	295	295	30	221
Hexatoma	230	400	295	295	25	236
Hydropsyche	296	400	423	423	400	409
Hydroptilidae	294	375	295	295	350	324
Lebertia	167	200	295	295	50	253
Lepidostoma	260	500	423	295	50	375
Leuctridae	153	200	295	295	150	221
Limnophila	211	300	295	295	50	324
Malenka	480	500	423	423	500	481
Micrasema	158	25	30	295	25	245
Narpus concolor	166	400	30	30	150	121

Table A.3 (cont.)

Nemata	291	375	295	295	25	324
Neophylax	225	225	201	295	225	324
Neoplasta	392	400	423	423	400	409
Oligochaeta	296	375	423	30	50	324
Optioservus	368	400	423	295	300	409
Optioservus quadrimaculatus	382	500	423	295	300	409
Paraleptophlebia	224	25	423	30	25	236.5
Pericoma	106	75	70	295	100	253
Perlodidae	141	100	70	30	75	153
Physa	374	375	423	423	375	409
Pisidium	363	425	423	423	30	409
Probezzia	119	25	30	30	50	121
Protzia	355	500	423	295	300	409
Rhithrogena	109	50	295	295	50	96
Rhyacophila	205	425	423	295	100	272
Rhyacophila brunnea vemna	193	325	295	295	50	253
Rhyacophila sibirica group B	200	50	423	30	75	394
Serratella	115	100	201	295	75	164
Serratella tibialis	107	75	70	295	50	201
Simulium	145	100	201	201	50	236
Sperchon	309	500	423	295	50	324
Suwallia	144	250	295	295	75	184
Sweltsa	157	200	295	295	25	209.8
Torrenticola	101	100	30	30	25	121
Turbellaria	265	200	295	295	25	316.8
Yoraperla	379	425	423	30	75	88.5
Zaitzevia	118	25	70	295	175	237
Zapada	170	400	30	295	25	192.5

Table A.4. Estimates of habitat preference of all 63 taxa for velocity (m/s) by six different methods. WA = weighted averaging, DPeak = peak of density curve, O:E = observed:expected ratios, SI = selectivity indices, PPeak = peak of the presence curve, and 75^{th} = 75^{th} percentiles.

Taxon	WA	DPeak	O:E	SI	PPeak	75th
Acentrella	0.36	0.4	0.4	1.33	0.3	0.4
Agapetus	0.25	0.4	0.25	0.25	0.25	0.33
Ameletus	0.25	0	0.14	0.67	0.3	0.4
Antocha monticola	0.48	0.2	0.67	1.33	0.25	0.67
Apatania	0.27	0.4	0.25	0.67	0.2	0.31
Arctopsyche grandis	0.58	0.7	0.67	1.33	0.2	0.6
Baetis	0.41	0.4	0.4	1.33	0.4	0.5
Brachycentrus americanus	0.31	0.2	0.14	1.33	0.2	0.59
Capniidae	0.20	0.1	0.14	0.14	0.2	0.4
Centroptilum	0.30	0.3	0.4	1.33	0.25	0.4
Cinygmula	0.45	0.2	1.33	1.33	0.25	0.5
Cleptelmis addenda	0.33	0.2	0.4	0.4	0.4	0.4
Dicosmoecus	0.31	0.2	0.67	0.67	0.2	0.4
Dicranota	0.34	0.4	0.4	1.33	0.2	0.4
Diphetor hageni	0.33	0.5	0.4	0.67	0.3	0.4
Drunella	0.32	0.4	0.25	1.33	0.2	0.4
Drunella coloradensis flavilinea	0.59	0.7	0.67	1.33	0.3	0.67
Drunella doddsii	0.34	0.1	0.25	1.33	0.25	0.5
Epeorus	0.53	1.3	1.33	1.33	0.5	0.55
Ephemerella inermis dorothea	0.31	0.1	0.4	0.67	0.35	0.5
Glossosoma	0.50	0.3	1.33	1.33	0.25	0.67
Glutops	0.47	0.5	1.33	1.33	0.3	0.67
Hesperoperla pacifica	0.36	0.4	0.4	0.67	0.2	0.5
Heterlimnius	0.42	0.2	0.4	1.33	0.3	0.45
Heterlimnius corpulentus	0.36	0.1	0.4	1.33	0.3	0.5
Hexatoma	0.29	0.3	0.14	0.67	0.25	0.4
Hydropsyche	0.44	0.2	1.33	1.33	0.45	0.5
Hydroptilidae	0.26	0.1	0.14	0.67	0.25	0.4
Lebertia	0.25	0.2	0.14	1.33	0.25	0.4
Lepidostoma	0.31	0.2	0.14	1.33	0.25	0.5
Leuctridae	0.59	0.6	1.33	1.33	0.4	0.67
Limnophila	0.22	0.3	0.14	0.14	0.25	0.33
Malenka	0.40	0.4	0.4	0.67	0.3	0.4
Micrasema	0.40	0.4	0.4	1.33	0.25	0.5
Narpus concolor	0.43	0.3	0.25	1.33	0.25	0.4

Table A.4 (cont.)

Nemata	0.28	0.1	0.14	0.67	0.3	0.45
Neophylax	0.45	0.2	0.67	1.33	0.75	0.67
Neoplasta	0.38	0.4	0.4	1.33	0.4	0.5
Oligochaeta	0.27	0.1	0.14	1.33	0.6	0.5
Optioservus	0.30	0.4	0.25	1.33	0.3	0.4
Optioservus quadrimaculatus	0.32	0.3	0.25	1.33	0.4	0.4
Paraleptophlebia	0.32	0.7	0.67	0.67	0.25	0.4
Pericoma	0.26	0.2	0.14	0.67	0.25	0.5
Perlodidae	0.32	0.1	0.14	1.33	0.25	0.4
Physa	0.20	0.2	0.14	0.14	0.25	0.32
Pisidium	0.19	0.1	0.14	0.67	0.3	0.4
Probezzia	0.37	0.3	1.33	1.33	0.3	0.4
Protzia	0.40	0.5	0.4	1.33	0.4	0.44
Rhithrogena	0.52	0.25	1.33	1.33	0.1	0.67
Rhyacophila	0.47	0.2	1.33	1.33	0.4	0.5
Rhyacophila brunnea vemna	0.43	0.4	0.4	1.33	0.4	0.5
Rhyacophila sibirica group B	0.35	0.1	0.67	0.67	0.3	0.5
Serratella	0.46	1.3	1.33	1.33	0.25	0.4
Serratella tibialis	0.37	0.3	0.4	1.33	0.15	0.5
Simulium	0.48	0.5	0.4	1.33	0.3	0.5
Sperchon	0.34	0.4	0.4	1.33	0.3	0.4
Suwallia	0.39	0.3	0.25	1.33	0.3	0.4
Sweltsa	0.33	0.2	0.14	0.67	0.2	0.43
Torrenticola	0.28	0.2	0.14	0.67	0.25	0.4
Turbellaria	0.36	0.3	0.4	1.33	0.3	0.43
Yoraperla	0.39	0.4	0.4	0.4	0.2	0.4
Zaitzevia	0.42	0.4	1.33	1.33	0.3	0.5
Zapada	0.32	0.3	0.4	1.33	0.1	0.4

Table A.5. Estimates of habitat preference of all 63 taxa for median particle size (mm) by six different methods. WA = weighted averaging, DPeak = peak of density curve, O:E = observed:expected ratios, SI = selectivity indices, PPeak = peak of the presence curve, and 75^{th} = 75^{th} percentiles.

Taxon	WA	DPeak	O:E	SI	PPeak	75th
Acentrella	52.0	55	75.6	75.6	40	62.9
Agapetus	35.1	30	41	75.6	40	53.8
Ameletus	32.6	10	6.7	6.7	20	58.5
Antocha monticola	48.3	15	75.6	75.6	40	64.9
Apatania	36.7	10	6.7	75.6	50	61
Arctopsyche grandis	58.2	55	75.6	75.6	75	68.1
Baetis	41.6	30	41	75.6	45	58.2
Brachycentrus americanus	49.0	45	41	75.6	40	59.4
Capniidae	27.7	5	6.7	6.7	30	54.3
Centroptilum	41.2	35	41	6.7	5	55.4
Cinygmula	43.6	80	41	75.6	40	57.4
Cleptelmis addenda	36.3	25	19.4	75.6	30	60.1
Dicosmoecus	43.4	25	75.6	75.6	35	68.1
Dicranota	21.1	5	6.7	6.7	55	57.3
Diphetor hageni	36.9	25	19.4	6.7	45	53.9
Drunella	26.1	5	6.7	6.7	50	51.9
Drunella coloradensis flavilinea	43.1	45	41	75.6	50	56.1
Drunella doddsii	52.4	45	58.8	75.6	50	64.3
Epeorus	53.5	50	58.8	75.6	50	60.5
Ephemerella inermis dorothea	31.0	25	19.4	75.6	25	54
Glossosoma	49.0	40	41	75.6	55	59.5
Glutops	36.2	5	6.7	6.7	60	62.5
Hesperoperla pacifica	46.8	55	41	6.7	40	59.9
Heterlimnius	34.0	45	6.7	6.7	10	61.1
Heterlimnius corpulentus	37.2	50	6.7	6.7	45	57.5
Hexatoma	28.9	15	19.4	6.7	30	50
Hydropsyche	40.7	80	75.6	75.6	40	54.2
Hydroptilidae	25.0	55	6.7	6.7	45	58.3
Lebertia	27.7	20	6.7	6.7	30	52.6
Lepidostoma	30.1	10	6.7	75.6	40	45
Leuctridae	38.1	40	19.4	58.8	40	54.4
Limnophila	20.9	10	6.7	6.7	10	35.4
Malenka	28.6	15	19.4	6.7	15	52.2
Micrasema	52.6	75	75.6	75.6	50	59.7
Narpus concolor	39.4	20	75.6	75.6	10	60.5

Table A.5 (cont.)

Nemata	21.9	10	6.7	75.6	50	56.5
Neophylax	45.2	70	58.8	6.7	5	56.2
Neoplasta	20.1	15	19.4	6.7	10	57.1
Oligochaeta	25.4	1	6.7	6.7	40	57.8
Optioservus	29.6	15	19.4	6.7	15	56
Optioservus quadrimaculatus	27.1	20	19.4	75.6	10	57.5
Paraleptophlebia	34.0	10	6.7	75.6	55	59
Pericoma	58.2	60	58.8	75.6	60	62
Perlodidae	43.3	20	75.6	75.6	55	59
Physa	21.8	1	6.7	75.6	25	55.9
Pisidium	24.9	25	6.7	6.7	10	52
Probezzia	35.7	35	6.7	6.7	10	56.4
Protzia	35.8	25	19.4	75.6	30	57.5
Rhithrogena	47.6	35	58.8	75.6	55	61.2
Rhyacophila	42.5	10	6.7	75.6	40	59.9
Rhyacophila brunnea vemna	45.6	55	58.8	75.6	40	59.5
Rhyacophila sibirica group B	37.1	45	6.7	6.7	55	60.7
Serratella	46.0	50	41	75.6	50	54.6
Serratella tibialis	55.7	50	58.8	75.6	60	65.6
Simulium	44.0	60	41	75.6	50	57
Sperchon	39.5	50	19.4	75.6	35	60.7
Suwallia	37.4	30	19.4	6.7	40	54.3
Sweltsa	38.9	45	58.8	6.7	40	59.6
Torrenticola	43.6	65	75.6	75.6	40	60.5
Turbellaria	42.2	40	75.6	75.6	35	57
Yoraperla	51.2	65	58.8	6.7	55	58.2
Zaitzevia	44.4	35	41	6.7	35	59
Zapada	41.9	50	75.6	75.6	55	61.2

Table A.6. Estimates of habitat preference of all 63 taxa for percent fines (%) by six different methods. WA = weighted averaging, DPeak = peak of density curve, O:E = observed:expected ratios, SI = selectivity indices, PPeak = peak of the presence curve, and 75^{th} = 75^{th} percentiles.

Taxon	WA	DPeak	O:E	SI	PPeak	75th
Acentrella	5.8	5	2.1	37.5	4	9.7
Agapetus	7.4	10	11.2	24.7	2	11.4
Ameletus	11.5	30	24.7	37.5	2	10.6
Antocha monticola	6.5	5	5.6	5.6	2	8.3
Apatania	7.1	15	11.2	5.6	2	6.8
Arctopsyche grandis	3.3	5	2.1	5.6	2	7.2
Baetis	7.4	10	11.2	37.5	2	10.6
Brachycentrus americanus	4.7	5	5.6	5.6	2	9.6
Capniidae	14.5	35	37.5	37.5	2	11
Centroptilum	7.5	5	5.6	37.5	2	13.2
Cinygmula	7.2	1	5.6	37.5	2	10
Cleptelmis addenda	8.1	10	11.2	24.7	2	11.4
Dicosmoecus	7.1	15	11.2	37.5	2	8.8
Dicranota	24.4	35	37.5	37.5	2	20.9
Diphetor hageni	7.6	1	2.1	37.5	2	11.4
Drunella	12.2	20	37.5	37.5	5	11.4
Drunella coloradensis flavilinea	9.8	15	11.2	37.5	2	11.4
Drunella doddsii	5.8	5	5.6	24.7	2	6.3
Epeorus	4.1	1	5.6	5.6	2	7
Ephemerella inermis dorothea	6.5	5	2.1	24.7	2	11.2
Glossosoma	6.0	5	5.6	5.6	2	8.3
Glutops	16.6	35	37.5	37.5	2	10.1
Hesperoperla pacifica	6.7	10	5.6	37.5	2	7.7
Heterlimnius	16.6	5	37.5	37.5	2	16.2
Heterlimnius corpulentus	9.9	5	11.2	37.5	2	12.1
Hexatoma	10.9	15	11.2	37.5	2	12.1
Hydropsyche	7.8	5	11.2	37.5	2	11.2
Hydroptilidae	45.9	95	37.5	37.5	2	11.3
Lebertia	13.2	10	11.2	37.5	2	12.4
Lepidostoma	12.5	25	24.7	24.7	5	10.6
Leuctridae	9.4	20	11.2	24.7	4	10.5
Limnophila	18.9	25	24.7	37.5	2	29.5
Malenka	11.1	20	24.7	37.5	4	17.8
Micrasema	8.8	5	2.1	37.5	2	9.5
Narpus concolor	12.7	30	37.5	37.5	5	18.8

Table A.6 (cont.)

Nemata	32.3	95	37.5	37.5	2	13.1
Neophylax	9.0	5	37.5	37.5	2	11.7
Neoplasta	16.0	15	11.2	37.5	2	15.7
Oligochaeta	34.1	95	37.5	37.5	2	13.6
Optioservus	10.2	10	24.7	37.5	2	13.2
Optioservus quadrimaculatus	10.8	10	24.7	37.5	2	11.8
Paraleptophlebia	8.3	15	11.2	37.5	2	8.7
Pericoma	4.0	1	2.1	11.2	5	13.4
Perlodidae	7.8	25	24.7	24.7	2	9.4
Physa	45.8	95	37.5	37.5	2	13.5
Pisidium	21.2	20	24.7	37.5	2	16.8
Probezzia	11.4	10	24.7	37.5	2	13.7
Protzia	6.4	1	2.1	24.7	2	8.3
Rhithrogena	5.9	10	2.1	5.6	5	6.4
Rhyacophila	10.7	35	37.5	37.5	2	10.6
Rhyacophila brunnea vemna	6.5	1	2.1	24.7	2	10.7
Rhyacophila sibirica group B	12.9	5	37.5	37.5	5	16.4
Serratella	6.0	5	5.6	37.5	2	9.6
Serratella tibialis	3.4	5	2.1	5.6	2	6.7
Simulium	7.4	10	11.2	37.5	2	11.7
Sperchon	6.2	10	2.1	11.2	2	8.3
Suwallia	8.7	20	11.2	37.5	2	10.1
Sweltsa	9.5	5	24.7	37.5	2	11
Torrenticola	8.2	5	24.7	24.7	2	10.5
Turbellaria	7.0	5	5.6	37.5	2	11.1
Yoraperla	13.3	10	37.5	37.5	2	14.3
Zaitzevia	8.9	5	5.6	37.5	2	12.8
Zapada	12.7	35	37.5	37.5	2	10.4

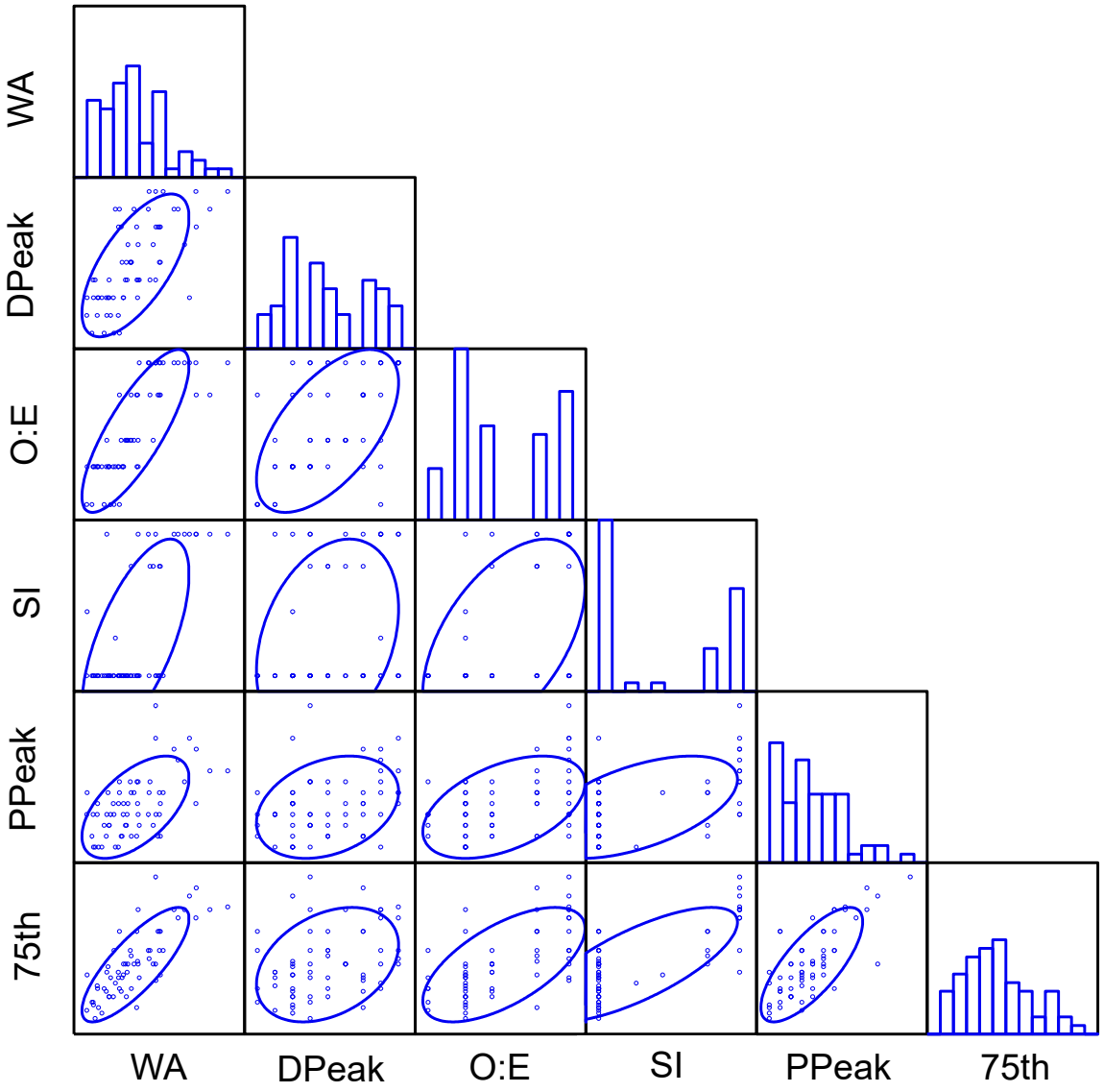

Figure A.1. Scatterplot matrix (SPLOM) showing relationships between the six types of estimates for temperature across all taxa. The SPLOM is shown with 75% normal distribution-based confidence ellipses. The * indicates that the relationship is significant at a Bonferroni-adjusted alpha of 0.003. WA = weighted averages, DPeak = peak of the density curve, O:E = observed:expected ratios, SI = selectivity indices, PPeak = peak of the presence curve, and 75^{th} = 75^{th} percentiles.

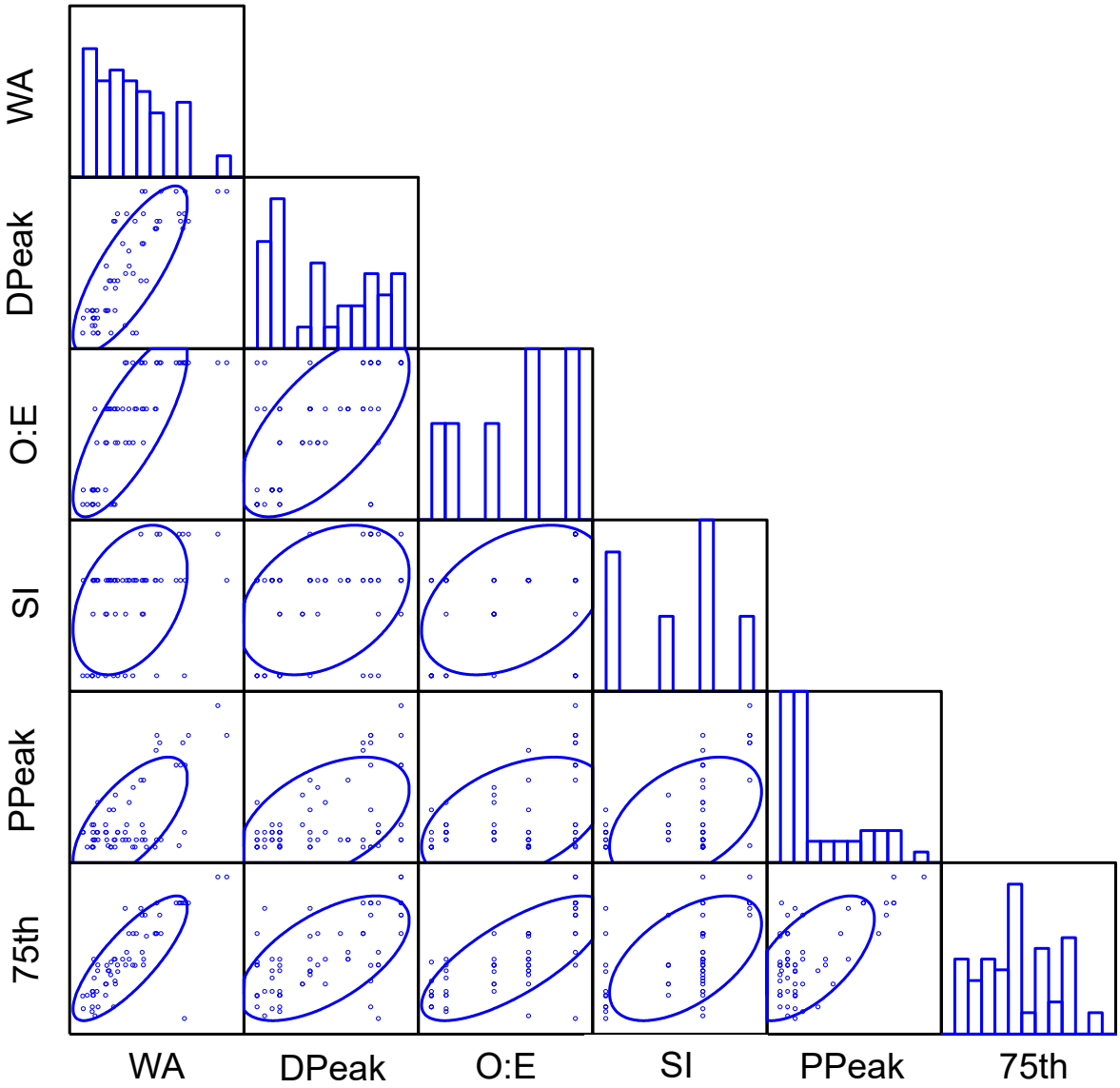

Figure A.2. Scatterplot matrix (SPLOM) showing relationships between the six types of estimates for conductivity across all taxa. The SPLOM is shown with 75% normal distribution-based confidence ellipses. The * indicates that the relationship is significant at a Bonferroni-adjusted alpha of 0.003. WA = weighted averages, DPeak = peak of the density curve, O:E = observed:expected ratios, SI = selectivity indices, PPeak = peak of the presence curve, and 75^{th} = 75^{th} percentiles.

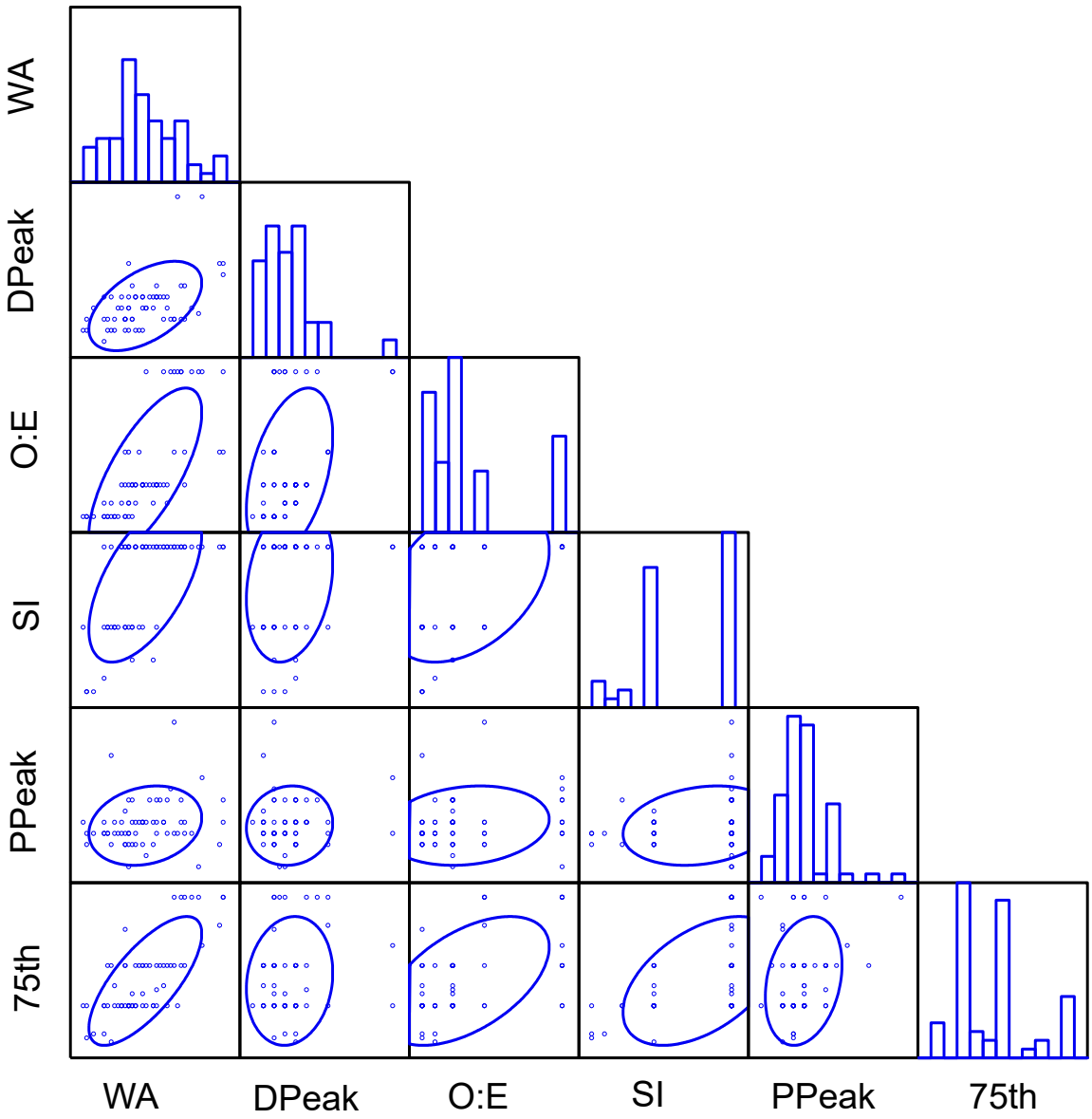

Figure A.3. Scatterplot matrix (SPLOM) showing relationships between the six types of estimates for velocity across all taxa. The SPLOM is shown with 75% normal distribution-based confidence ellipses. The * indicates that the relationship is significant at a Bonferroni-adjusted alpha of 0.003. WA = weighted averages, DPeak = peak of the density curve, O:E = observed:expected ratios, SI = selectivity indices, PPeak = peak of the presence curve, and 75^{th} = 75^{th} percentiles.

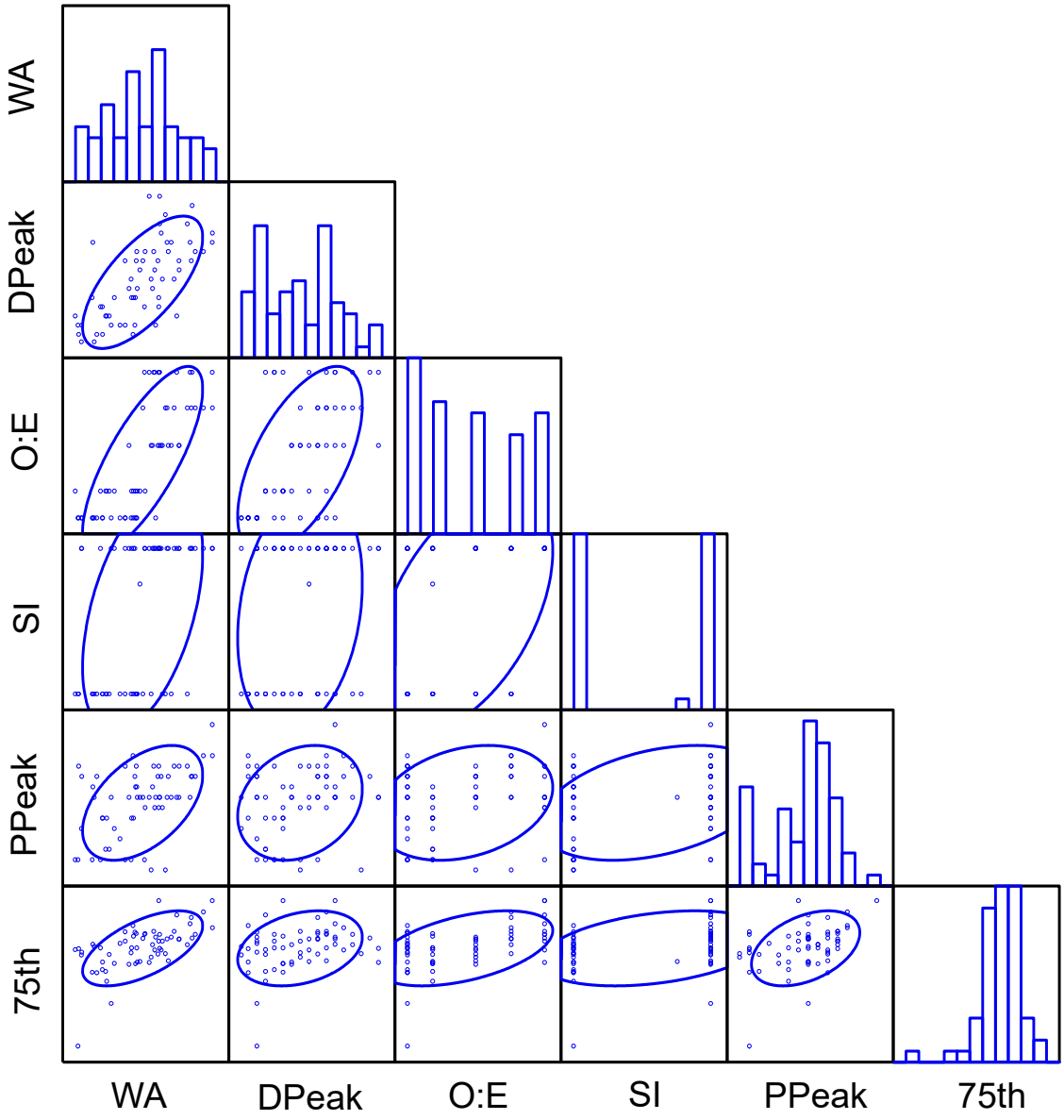

Figure A.4. Scatterplot matrix (SPLOM) showing relationships between the six types of estimates for median particle size across all taxa. The SPLOM is shown with 75% normal distribution-based confidence ellipses. The * indicates that the relationship is significant at a Bonferroni-adjusted alpha of 0.003. WA = weighted averages, DPeak = peak of the density curve, O:E = observed:expected ratios, SI = selectivity indices, PPeak = peak of the presence curve, and 75^{th} = 75^{th} percentiles.

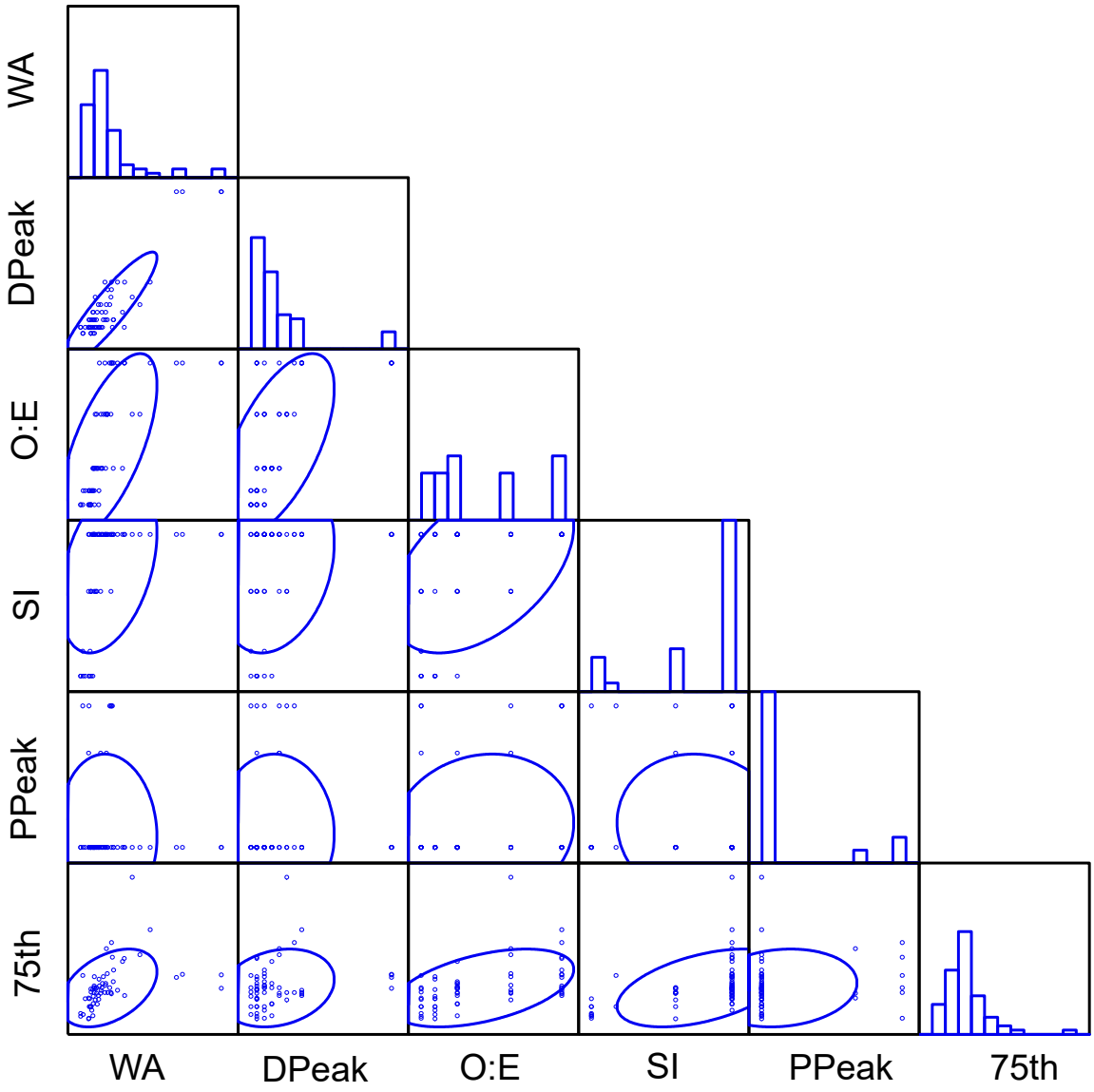

Figure A.5. Scatterplot matrix (SPLOM) showing relationships between the six types of estimates for percent fines across all taxa. The SPLOM is shown with 75% normal distribution-based confidence ellipses. The * indicates that the relationship is significant at a Bonferroni-adjusted alpha of 0.003. WA = weighted averages, DPeak = peak of the density curve, O:E = observed:expected ratios, SI = selectivity indices, PPeak = peak of the presence curve, and 75^{th} = 75^{th} percentiles.